

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:47:46 : Search time 33.11 Seconds
(Without alignments)
1583.414 Million cell updates/sec

Title: US-09-841-758-1
Perfect score: 472
Sequence: 1 MATKCGNCGYSTPLEAMK.....PALAHLRYPGDCSSDIWI 472

Scoring table: OLIGO
Gapop 60.0, Gapept 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_032802.*
1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	100.0	472	19 AAW37940	Amino acid sequence
2	472	100.0	472	21 AAY68328	Amyotrophic latera
3	457	96.8	499	21 AAB57139	Human prostate can
4	217	46.0	235	22 AAB95579	Human protein sequ
5	102	21.6	582	22 AAU30396	Novel human secret
6	77	16.3	138	21 AAB54427	Human pancreatic c
7	20	4.4	21	21 AAY68338	Amyotrophic latera
8	20	4.2	20	21 AAY68332	Amyotrophic latera
9	18	3.8	93	22 AA008400	Human polypeptide
10	17	3.6	17	21 AAY68335	Amyotrophic latera
11	16	3.4	21	AAV68333	Amyotrophic latera

12	16	3.4	167	21 AAG19618	Arabidopsis thalia
13	16	3.4	188	21 AAG19617	Arabidopsis thalia
14	16	3.4	199	21 AAG25100	Arabidopsis thalia
15	16	3.4	199	21 AAG38106	Arabidopsis thalia
16	16	3.4	236	21 AAG16911	Arabidopsis thalia
17	16	3.4	236	21 AAG25099	Arabidopsis thalia
18	16	3.4	275	21 AAG16910	Arabidopsis thalia
19	16	3.4	275	21 AAG25098	Arabidopsis thalia
20	16	3.4	275	21 AAG38105	Arabidopsis thalia
21	16	3.4	327	21 AAG38104	Arabidopsis thalia
22	16	3.4	328	21 AAG16909	Arabidopsis thalia
23	16	3.4	409	21 AAG36720	Arabidopsis thalia
24	16	3.4	455	21 AAG17316	Arabidopsis thalia
25	16	3.4	455	21 AAG36719	Arabidopsis thalia
26	16	3.4	488	21 AAG17315	Arabidopsis thalia
27	16	3.4	490	21 AAG36718	Arabidopsis thalia
28	16	3.4	503	21 AAG17314	Arabidopsis thalia
29	15	3.2	59	22 AAG30395	Novel human secret
30	15	3.2	409	21 AAG19253	Arabidopsis thalia
31	15	3.2	469	21 AAG19252	Arabidopsis thalia
32	15	3.2	502	21 AAG19251	Arabidopsis thalia
33	14	3.0	14	21 AAY68337	Amyotrophic latera
34	12	2.5	92	22 AAU32310	Novel human secret
35	11	2.3	11	21 AAY68334	Amyotrophic latera
36	10	2.1	10	21 AAY68330	Amyotrophic latera
37	10	2.1	455	21 AAG52478	Arabidopsis thalia
38	10	2.1	468	22 ABB71631	Drosophila melanog
39	10	2.1	476	21 AAG52477	Arabidopsis thalia
40	10	2.1	480	21 AAG52476	Arabidopsis thalia
41	9	1.9	9	21 AAY68331	Amyotrophic latera
42	8	1.7	118	22 AAM88926	Human immune/haema
43	8	1.7	299	22 AAU50496	Propionibacterium
44	7	1.5	9	22 AAG98228	Human SNP associat
45	7	1.5	58	22 AAG61652	Propionibacterium

ALIGNMENTS

RESULT 1	
ID AAW37940	standard; Protein; 472 AA.
XX	
AC AAW37940:	
XX	
DT 10-AUG-1998	(first entry)
XX	
DE	Amino acid sequence of the human selenium-binding protein.
KW	Human selenium-binding protein; HSEBP; recombinant HSEBP; liver damage;
KM	kidney damage; lung damage; antibody; cancer.
XX	
OS	Homo sapiens.
XX	
PN	US5759812-A.
XX	
PD	02-JUN-1998.
XX	
PF	15-NOV-1996; 96US-0749903.
XX	
PR	15-NOV-1996; 96US-0749903.
XX	
PA	(INCY-) INCYTE PHARM INC.
XX	
PI	Bandman O, Hawkins PR;
XX	
DR	WPI; 1998-332133/29.
XX	
PT	N-PDB; AAV29245.
XX	
PS	DNA encoding human selenium-binding protein - useful for producing recombinant protein
XX	
PS	Claim 1: Columns 33-36; 35pp; English.

XX This is the amino acid sequence of the novel human selenium-binding
 CC protein (HSEBP). In the method of the invention, recombinant HSEBP
 CC is produced which can be used to treat liver, kidney or lung damage
 CC caused by chemical agents, or to produce antibodies. The vector can
 CC be used to treat liver, kidney or lung damage caused by chemical
 CC agents or to treat cancer.

XX
 SQ Sequence 472 AA:

Query Match 100.0%; Score 472; DB 19; Length 472;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MATKGCNGCPGYSTPLEAMKGPREEIYVLPICIRNTGTEADPYLATVDVDPKSPQCYVI 60
 |||||||
 Db 1 matkcgngcpgystpleamkgpreeivlylpcirntgteapdylatvdvdpkspqcyvi 60
 OY 61 HRLPMPNLKDELHHSQWNTCSGCFDSTKSRKLVLPSSLISRIYVVDVGSPPRAPKLHK 120
 |||||||
 Db 61 hrlpmpnlkdelhhsqwtncscfsgdstksrklvlpsslissriyvvdvgspprapklhk 120
 OY 121 VIEPDIHAKCELAFLHSHCLASGEVMISSLDYKGNKGKGFVLIDGTEFEVKGTEWRP 180
 |||||||
 Db 121 viepdihaaceiaflhshclasegmisldgkngkgfvlldgetfevkgtewrp 180
 OY 181 GGAAPLGVDWFYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYSHLYWDMORHEIVQ 240
 |||||||
 Db 181 ggaaplgvdfwyqprhnwmiSTEMAAPNVLRDGFNPADVEAGLYSHLYWDMORHEIVQ 240
 OY 241 TSLKDGILPIELIRFLHNPDAAGFVGALSTIQRFKNEGQTSVEKVIQVPPKKYKG 300
 |||||||
 Db 241 tslkdgllpielirflhnpdaagfvgalsstiqrfkneqgtsvekviqvpkkkyg 300
 OY 301 WLLPMPGLITPILSLDRLFYFSNWLHGDLYOYDISDPPORRLTGOLFSGSIVKGGP 360
 |||||||
 Db 301 wllpmpglitdillsldrflfysnwlhgdlyoydisdpporrltgoelfsgsivkggp 360
 OY 361 VOYLEDEELKSQPEPLVYKGRVAGGPMQISLDGKRLYITTSLSXAMDQFYDILRE 420
 |||||||
 Db 361 vglyedeelksqpeplvykgrvaggpqmqlsldgkrllyttslysawdkfydpdlire 420
 OY 421 GSVMLQVVDVDFYKGLKLNPNFLVDFGKEPLGPAALAEHLRYGGGSSDIWI 472
 |||||||
 Db 421 gsvmlqvvdvdfykglklpnflvdfgkeplgpalahelrpyggdssdiwi 472

RESULT 2

ID AAY68328 standard; Protein: 472 AA.

XX AAY68328;

DT 17-APR-2000 (first entry)

DE Amyotrophic lateral sclerosis related p53 protein.

KW Amyotrophic lateral sclerosis; p53; ALS; diagnosis; familial ALS.

XX Homo sapiens.

PN JP2000000095-A.

PD 07-JAN-2000.

PF 15-JUN-1998; 98JP-0167259.

PR 15-JUN-1998; 98JP-0167259.

PA (SAKO/) SAKODA S.

XX (IGAK-) IGAKU SEIBUNSUOGAKU KENKYUSHO KK.

DR WPI: 2000-129429/12.

DR N-PSDB: AAZ57589.

PT New DNA encoding the p53 protein - useful in the diagnosis of

PS amyotrophic lateral sclerosis

CC Claim 1; Page 6-8; 11pp; Japanese.

CC The present sequence represents the protein p53 which is related to

CC amyotrophic lateral sclerosis (ALS). The protein p53 and its gene

CC can be used for the diagnosis of ALS.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

PD 21-SEP-2000.
 XX
 PF 08-MAR-2000; 2000MO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2000-587513/55.
 DR N-PSDB; AAF16342.
 XX
 XX Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 PS Claim 11; Page 2190-2192; 2338pp; English.

CC AAF15566-0 AAF15505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB55363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytoskeletal,
CC cardioactive, immunomodulatory, muscular, vlinerary, gastrointestinal,
CC neptrotoxic, antihelictive, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF15506 to AAF15514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.

xx
xx

SQ Sequence 499 AA;

Query Match 96.8%; Score 457; DB 21; Length 499;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 457; Conservative 0; Mismatches 0; Indels 0; Gaps 0

RESULT	4
AAB95579	
ID	AAB95579 standard; Protein; 235 AA.
XX	
AC	AAB95579;
XX	
DT	26-JUN-2001 (first entry)
XX	
DE	Human protein sequence SEQ ID NO:18237.
XX	
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX	
OS	Homo sapiens.
XX	
PN	EP1074617-A2.
XX	
PD	07-FEB-2001.
XX	
PE	28-JUL-2000; 2000EP-0116126.
XX	
PR	29-JUL-1999; 99JP-0248036.
PR	27-AUG-1999; 99JP-0300253.
PR	11-JAN-2000; 2000JP-0118776.
PR	02-MAY-2000; 2000JP-0183767.
PR	09-JUN-2000; 2000JP-0241899.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX	
DR	WPI; 2001-318749/34.
XX	
PT	Primer sets for synthesizing polynucleotides, particularly the 5602
PT	full-length cDNAs defined in the specification, and for the detection
PT	and/or diagnosis of the abnormality of the proteins encoded by the
PT	full-length cDNAs -
XX	
XX	Claim 8; SEQ ID 18237; 2537pp + CD ROM; English.

Db	388	vqviedeelkqgpeplvvxkygrvagvgpmqlslidckrllyttslysawdkfyppalire	447
Qy	421	GSYMLQVDVDTVKGGLKLNPNFLVDFEGKEPIGPLALH	457
Db	448	gsyamlqvadvtkggjlklnpnflvdftgkepi GPLAH	464

Qy	65	MPNLDELHHSGWNTCCSCFCDSTKSRTKLVLPSLISSRIYYDVDCSEPPAPLHLHVIEP	124
----	----	---	-----

Query Match 46.0%; Score 217; DB 22; Length 225;

Best Local Similarity 100.0%; Pred. No. 2, 6e-208;

Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

|||||
Db 1 mpnldehshgwnccsfsgdstskrtklylpsllssriyvvdygseprapklkvleap 60
OY 125 KOIHAKCELAFHTSHCLASGEWMISSLGDKVNGKGGVLLDGFEPFKGWBRPGGAA 184
Db 61 kdihtkcelafhtshchlasgewmissslgdvkngkgfvlldgetfevkgtwepgpgaa 120
OY 185 PLGYDFWYQPRHNWISTEMAAPNVLRDGFNPADVEAGLYGSHLYVMDQMRHEIVQTLSTL 244
Db 121 ptygdfwyqprhnwisteweaapnvlrdgfnpadveaglygshlyvwdwqrheivqtlsl 180
OY 245 KDGLPLFLFRFHNPDAAAGFYGCALSSSTIOREYKNE 281
Db 181 kdglplflfrfhnppdaagfyvgcalssctiqrykne 217

RESULT 5
AAU30396
ID AAU30396 standard; Protein: 582 AA.
XX
AC AAU30396;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #887.
XX
KW Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200179449-A2.
XX
PD 25-OCT-2001.
XX
PE 16-APR-2001; 2001WO-US08656.
XX
PR 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-611725/70.
XX
PT Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy -
PS Claim 20; Page 288; 765pp; English.

XX
XX The invention relates to novel human secreted polypeptides. The
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression or
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
XX sequences of novel human secreted proteins of the invention.

Sequence 582 AA;

Query Match 21.6%; Score 102; DB 22; Length 582;
Best Local Similarity 100.0%; Pred. No. 5.3e-93;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 134 AFLHTSHCLASGEWMISSLGDKVNGKGGFVYLLOSETFEVKGWBRPGAAPLGYDFWYQ 193
Db 175 aflhtshclasewmissslgdvkngkgfvlldgetfevkgtwepgpgaa1gydfwyq 234
OY 194 PRHNWISTEMAAPNVLRDGFNPADVEAGLYGSHLYVMDQMR 235
Db 235 prhnwisteweaapnvlrdgfnpadveaglygshlyvwdwqr 276

RESULT 6
AAB54427
ID AAB54427 standard; Protein: 138 AA.
XX
AC AAB54427;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human pancreatic cancer antigen protein sequence SEQ ID NO:879.
XX
KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW neotropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiac; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative.
XX
XX Homo sapiens.
XX
OS
XX
PN WO200055320-A1.
XX
PD 21-SEP-2000.
XX
PE 08-MAR-2000; 2000WO-US05989.
XX
PR 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-579444/54.
DR N-PSDB; AAC99192.
XX
PT New nucleic acid that is a pancreatic cancer antigen for preventing,
XX treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition -
PS Claim 11; Page 1339; 1379pp; English.

XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX proteins, called pancreatic cancer antigens, given in AAB54008 to
XX AAB54466. The human pancreatic cancer antigens have cytostatic,
XX neuroprotective, neotropic, immunomodulatory, relaxant, contraceptive,
XX gynaecological, cardiac and antiinflammatory activities, and can be used
XX in gene therapy. The polynucleotide and proteins can be used for
XX preventing, treating, or ameliorating a medical condition or in assays
XX for diagnosing a pathological condition or a susceptibility to one in a
XX subject. Binding partners to the proteins and the activity of the
XX proteins can be identified. The pancreatic cancer antigens can be used to
XX detect, treat or prevent pancreatic disorders, especially cancer.
XX Agonists and antagonists to the antigens can be screened for. The
XX pancreatic cancer antigen polynucleotides can be used to design nucleic
XX acid hybridisation probes that can be used in chromosome mapping, linkage
XX analysis, tissue identification and/or typing and a variety of forensic
XX and diagnostic methods. The proteins can be used to generate antibodies
XX which are used to purify, detect and target the polypeptides, including
XX both in vivo and in vitro diagnostic and therapeutic methods. The
XX proteins can be used to treat or prevent neural, immune system, muscular,

CC reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
 CC proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
 CC sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 138 AA;

Query Match 16.3%; Score 77; DB 21; Length 138;
 Best Local Similarity 100.0%; Pred. No. 1.4e-68;
 Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 132 ELAFLHTSHCLAGSEWITSLGDKGKNGKGFVLLDGETFEVKGTEWRPGAPLGIDFW 191
 |||||||
 Db 6 elafhlshclagsewmislgdkngkgfvlldgetfevkgtewrpgaaipydfw 65

OY 192 YOPRHNVNISTENAAPN 208
 |||||||
 Db 66 yqprhnmvstewaapn 82

RESULT 7

AAV68338
 ID AAV68338 standard; peptide; 21 AA.
 XX
 AC AAV68338;

DT 17-APR-2000 (first entry)

XX Amnrotrophic lateral sclerosis related p53 protein peptide #9.

XX Amnrotrophic lateral sclerosis; p53; ALS; diagnosis; familial ALS.

XX Homo sapiens.

XX JP2000000095-A.

PD 07-JAN-2000.

PF 15-JUN-1998; 98JP-0167259.

PR 15-JUN-1998; 98JP-0167259.

PA (SAKO/) SAKODA S.
 PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.

DR WPI; 2000-129429/12.

PT New DNA encoding the p53 protein - useful in the diagnosis of

PT amnrotrophic lateral sclerosis

PS Disclosure; Page 5; 11pp; Japanese.

XX The present sequence represents a peptide from the protein p53 which
 CC is related to amnrotrophic lateral sclerosis (ALS). The protein p53 and
 CC its gene can be used for the diagnosis of ALS.
 XX

SQ Sequence 21 AA;

Query Match 4.4%; Score 21; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 2.3e-13;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 69 KDELHSGMNTCSGCFDSTK 89
 |||||||
 Db 1 kdelhsgmntcsfcgdstck 21

RESULT 8

AAV68332
 ID AAV68332 standard; peptide; 20 AA.
 XX
 AC AAV68332;

XX 17-APR-2000 (first entry)

XX Amnrotrophic lateral sclerosis related p53 protein peptide #4.

XX Amnrotrophic lateral sclerosis; p53; ALS; diagnosis; familial ALS.

XX Homo sapiens.

XX JP2000000095-A.

PD 07-JAN-2000.

PF 15-JUN-1998; 98JP-0167259.

PR 15-JUN-1998; 98JP-0167259.

PA (SAKO/) SAKODA S.
 PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.

DR WPI; 2000-129429/12.

PT New DNA encoding the p53 protein - useful in the diagnosis of

PT amnrotrophic lateral sclerosis

PS Disclosure; Page 5; 11pp; Japanese.

XX The present sequence represents a peptide from the protein p53 which
 CC is related to amnrotrophic lateral sclerosis (ALS). The protein p53 and
 CC its gene can be used for the diagnosis of ALS.
 XX

SQ Sequence 20 AA;

Query Match 4.2%; Score 20; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 2.2e-12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 DELHSGMNTCSGCFDSTK 89
 |||||||
 Db 1 delhsgmntcsfcgdstck 20

RESULT 9

AAO08400
 ID AAO08400 standard; Protein; 93 AA.
 XX
 AC AAO08400;

DT 06-NOV-2001 (first entry).

DE Human polypeptide SEQ ID NO 22292.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT.

DR WPI: 2001-514838/56.
 DR N-PSDB; AA188331.
 PT Isolated nucleic acids and polypeptides, useful for preventing
 PT diagnosing and treating e.g. leukemia, inflammation and immune
 PT disorders -
 XX
 PS Claim 20; SEQ ID NO 22292; 1399pp + Sequence Listing; English.
 XX
 CC The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA000010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 CC
 SQ Sequence 93 AA;

Query Match 3.8%; Score 18; DB 22; Length 93;
 Best Local Similarity 100.0%; Pred. No. 9,1e-10;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 455 LAHELRYPGDCSSDIWI 472
 ||||||||||||||||
 DB 76 lahelrypgdcssdiwi 93

RESULT 10
 AAY68335
 ID AAY68335 standard; peptide; 17 AA.
 XX
 AC AAY68335;
 XX
 DT 17-APR-2000 (first entry)
 XX
 DE Amyotrophic lateral sclerosis related p53 protein peptide #7.
 XX
 KW Amyotrophic lateral sclerosis; p53; ALS; diagnosis; familial ALS.
 XX
 OS Homo sapiens.
 XX
 PN JP2000000095-A.
 XX
 PD 07-JAN-2000.
 XX
 PF 15-JUN-1998; 98JP-0167259.
 XX
 PR 15-JUN-1998; 98JP-0167259.
 XX
 PA (SAKO/) SAKODA S.
 PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.
 XX
 DR WPI: 2000-129429/12.
 XX
 PT New DNA encoding the p53 protein - useful in the diagnosis of
 PT amyotrophic lateral sclerosis
 XX
 PS Disclosure; Page 5; 11pp; Japanese.
 XX
 CC The present sequence represents a peptide from the protein p53 which
 CC is related to amyotrophic lateral sclerosis (ALS). The protein p53 and
 CC its gene can be used for the diagnosis of ALS.
 CC
 SQ Sequence 17 AA;

Query Match 3.6%; Score 17; DB 21; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1,9e-09;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 412 OFYPLIREGSVMLOVD 428
 ||||||||||||||||
 DB 1 qfypdliregsvmlqvd 17

RESULT 11
 AAY68333
 ID AAY68333 standard; peptide; 16 AA.
 XX
 AC AAY68333;
 XX
 DT 17-APR-2000 (first entry)
 XX
 DE Amyotrophic lateral sclerosis related p53 protein peptide #5.
 XX
 KW Amyotrophic lateral sclerosis; p53; ALS; diagnosis; familial ALS.
 XX
 OS Homo sapiens.
 XX
 PN JP2000000095-A.
 XX
 PD 07-JAN-2000.
 XX
 PF 15-JUN-1998; 98JP-0167259.
 XX
 PR 15-JUN-1998; 98JP-0167259.
 XX
 PA (SAKO/) SAKODA S.
 PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.
 XX
 DR WPI: 2000-129429/12.
 XX
 PT New DNA encoding the p53 protein - useful in the diagnosis of
 PT amyotrophic lateral sclerosis
 XX
 PS Disclosure; Page 5; 11pp; Japanese.
 XX
 CC The present sequence represents a peptide from the protein p53 which
 CC is related to amyotrophic lateral sclerosis (ALS). The protein p53 and
 CC its gene can be used for the diagnosis of ALS.
 CC
 SQ Sequence 16 AA;

Query Match 3.4%; Score 16; DB 21; Length 16;
 Best Local Similarity 100.0%; Pred. No. 1,8e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 CGNCGPGYSTPLEAMK 20
 ||||||||||||||||
 DB 1 cgnccpgystpleamk 16

RESULT 12
 AAG19618
 ID AAG19618 standard; Protein; 167 AA.
 XX
 AC AAG19618;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 21487.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX

OS Arabidopsis thaliana.
 XX
 PN EPI033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 07-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134321.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 17-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143524.
 PR 15-JUL-1999; 99US-0143524.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144684.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145088.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145213.
 PR 27-JUL-1999; 99US-0145218.
 PR 27-JUL-1999; 99US-0145218.
 PR 27-JUL-1999; 99US-0145219.
 PR 28-JUL-1999; 99US-0145591.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 990S-0154018.
 PR 16-SEP-1999; 990S-0154039.
 PR 20-SEP-1999; 990S-0154779.
 PR 22-SEP-1999; 990S-0155139.
 PR 23-SEP-1999; 990S-0155486.
 PR 24-SEP-1999; 990S-0155659.
 PR 28-SEP-1999; 990S-0156458.
 PR 29-SEP-1999; 990S-0156596.
 PR 04-OCT-1999; 990S-0157117.
 PR 05-OCT-1999; 990S-0157753.
 PR 06-OCT-1999; 990S-0157865.
 PR 07-OCT-1999; 990S-0158029.
 PR 08-OCT-1999; 990S-0158232.
 PR 12-OCT-1999; 990S-0158369.
 PR 13-OCT-1999; 990S-0159293.
 PR 13-OCT-1999; 990S-0159294.
 PR 13-OCT-1999; 990S-0159295.
 PR 14-OCT-1999; 990S-0159329.
 PR 14-OCT-1999; 990S-0159330.
 PR 14-OCT-1999; 990S-0159331.
 PR 14-OCT-1999; 990S-0159637.
 PR 14-OCT-1999; 990S-0159638.
 PR 18-OCT-1999; 990S-0159584.
 PR 21-OCT-1999; 990S-0160741.
 PR 21-OCT-1999; 990S-0160767.
 PR 21-OCT-1999; 990S-0160768.
 PR 21-OCT-1999; 990S-0160770.
 PR 21-OCT-1999; 990S-0160814.
 PR 21-OCT-1999; 990S-0160815.
 PR 22-OCT-1999; 990S-0160980.
 PR 22-OCT-1999; 990S-0160981.
 PR 22-OCT-1999; 990S-0160989.
 PR 25-OCT-1999; 990S-0161404.
 PR 25-OCT-1999; 990S-0161405.
 PR 25-OCT-1999; 990S-0161406.
 PR 26-OCT-1999; 990S-0161359.
 PR 26-OCT-1999; 990S-0161360.
 PR 26-OCT-1999; 990S-0161361.
 PR 28-OCT-1999; 990S-0161920.
 PR 28-OCT-1999; 990S-0161922.
 PR 28-OCT-1999; 990S-0161993.
 PR 29-OCT-1999; 990S-0162142.

Query Match 3.4%; Score 16; DB 21; Length 167;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 385 GGPOMTOLSLDGKRLY 400
 Db 80 ggpqmqlslsgkrlly 95

RESULT 13
 AAG19617
 ID AAG19617 standard; Protein; 188 AA.
 AC AAG19617;
 DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 21486.
 XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 OS Arabidopsis thaliana.
 XX EPI033405-A2.
 PN EPI033405-A2.
 XX 06-SEP-2000.
 PD 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.
 PR 05-MAR-1999; 990S-0123180.
 PR 09-MAR-1999; 990S-0123548.
 PR 23-MAR-1999; 990S-0125788.
 PR 25-MAR-1999; 990S-0126264.
 PR 29-MAR-1999; 990S-0126785.
 PR 01-APR-1999; 990S-0127462.
 PR 06-APR-1999; 990S-0128234.
 PR 08-APR-1999; 990S-0128714.
 PR 16-APR-1999; 990S-0129845.
 PR 19-APR-1999; 990S-0130077.
 PR 21-APR-1999; 990S-0130449.
 PR 23-APR-1999; 990S-0130510.
 PR 23-APR-1999; 990S-0130891.
 PR 28-APR-1999; 990S-0131449.
 PR 30-APR-1999; 990S-0132048.
 PR 30-APR-1999; 990S-0132407.
 PR 04-MAY-1999; 990S-0132484.
 PR 05-MAY-1999; 990S-0132485.
 PR 06-MAY-1999; 990S-0132486.
 PR 06-MAY-1999; 990S-0132487.
 PR 07-MAY-1999; 990S-0132863.
 PR 11-MAY-1999; 990S-0134256.
 PR 14-MAY-1999; 990S-0134218.
 PR 14-MAY-1999; 990S-0134219.
 PR 14-MAY-1999; 990S-0134221.
 PR 14-MAY-1999; 990S-0134370.
 PR 18-MAY-1999; 990S-0134768.
 PR 19-MAY-1999; 990S-0134941.
 PR 20-MAY-1999; 990S-0135124.
 PR 21-MAY-1999; 990S-0135353.
 PR 25-MAY-1999; 990S-0135629.
 PR 27-MAY-1999; 990S-0136021.
 PR 28-MAY-1999; 990S-0136392.
 PR 01-JUN-1999; 990S-0137222.
 PR 03-JUN-1999; 990S-0137528.
 PR 04-JUN-1999; 990S-0137502.
 PR 07-JUN-1999; 990S-0137724.
 PR 08-JUN-1999; 990S-0138094.
 PR 10-JUN-1999; 990S-0138540.
 PR 10-JUN-1999; 990S-0138847.
 PR 14-JUN-1999; 990S-0139119.
 PR 16-JUN-1999; 990S-0139452.
 PR 16-JUN-1999; 990S-0139453.
 PR 17-JUN-1999; 990S-0139492.
 PR 18-JUN-1999; 990S-0139454.
 PR 18-JUN-1999; 990S-0139455.
 PR 18-JUN-1999; 990S-0139456.
 PR 18-JUN-1999; 990S-0139457.
 PR 18-JUN-1999; 990S-0139458.
 PR 18-JUN-1999; 990S-0139459.
 PR 18-JUN-1999; 990S-0139460.
 PR 18-JUN-1999; 990S-0139461.
 PR 18-JUN-1999; 990S-0139462.
 PR 18-JUN-1999; 990S-0139463.
 PR 18-JUN-1999; 990S-0139750.
 PR 18-JUN-1999; 990S-0139763.
 PR 21-JUN-1999; 990S-0139817.
 PR 22-JUN-1999; 990S-0139899.
 PR 23-JUN-1999; 990S-0140353.
 PR 23-JUN-1999; 990S-0140354.
 PR 24-JUN-1999; 990S-0140695.
 PR 28-JUN-1999; 990S-0140823.
 PR 29-JUN-1999; 990S-0140991.
 PR 30-JUN-1999; 990S-0141287.
 PR 01-JUL-1999; 990S-0141842.
 PR 01-JUL-1999; 990S-0142154.
 PR 02-JUL-1999; 990S-0142055.
 PR 06-JUL-1999; 990S-0142390.
 PR 08-JUL-1999; 990S-0142803.
 PR 09-JUL-1999; 990S-0142920.

```

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145216.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

```

```

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.4%; Score 16; DB 21; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 GGPMTQLSDGKRLY 400
    |||||
Db 101 ggpmtqlsdgkrlY 116

RESULT 14
AAG25100
ID AAG25100 standard; Protein; 199 AA.
XX
AC AAG25100;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 29026.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.

```

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130444.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145912.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.

PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 3.4%; Score 16; DB 21; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 385 GCPQMTLSIDGKRLY 400
|||||
Db 112 gspqmtlsgldgkrlly 127

RESULT 15

AAG38106
ID AAG38106 standard; Protein; 199 AA.

AC AAG38106;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 46960.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

OS Arabidopsis thaliana.

PN EPI033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0126264.
PR 29-MAR-1999; 990S-0126785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.

PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 07-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 11-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 18-MAY-1999; 990S-0134768.
PR 19-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 21-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 10-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139753.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 28-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142820.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148341.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 3.4%; Score 16; DB 21; Length 199;
Best Local Similarity 100.0%; Pred.No. 1.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 385 GGPOMIQSLDGKRLY 400
|||||
Db 112 gppmqlsldgkrlx 127

Search completed: August 29, 2002, 16:52:01
Job time: 255 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:49:46 ; Search time 16.35 Seconds
(Without alignments)
705.130 Million cell updates/sec

Title: US-09-841-758-1
Perfect score: 472
Sequence: 1 MATKCGNCGPXYRPLEAMK.....PALAHLRYPGDCCSSDIWI 472

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	472	100.0	472	1	US-08-749-903-1
2	472	100.0	472	4	US-09-088-641-1
3	124	26.3	472	1	US-08-749-903-3
4	124	26.3	472	4	US-09-088-641-3
5	49	10.4	472	1	US-08-749-903-4
6	49	10.4	472	4	US-09-088-641-4
7	44	9.3	472	1	US-08-749-903-5
8	44	9.3	472	4	US-09-088-641-5
9	7	1.5	472	1	US-08-858-207A-448
10	7	1.5	467	1	US-08-459-100A-3
11	7	1.5	500	1	US-08-307-499-20
12	7	1.5	500	4	US-09-299-268-20
13	7	1.5	685	3	US-09-031-563-21
14	7	1.5	788	2	US-08-907-166-6
15	7	1.5	1055	3	US-09-031-563-27
16	7	1.5	1315	3	US-09-031-563-2
17	7	1.5	1315	3	US-09-031-563-25
18	7	1.5	1315	4	US-09-293-505-10
19	6	1.3	18	4	US-09-261-855-30
20	6	1.3	18	4	US-08-468-514-5
21	6	1.3	64	2	US-08-687-559-3
22	6	1.3	69	2	US-08-726-306A-53
23	6	1.3	76	1	US-08-519-777-22
24	6	1.3	76	1	US-08-742-035-22
25	6	1.3	76	2	US-08-777-019-22
26	6	1.3	76	2	US-08-777-143-22
27	6	1.3	76	3	US-08-775-414-22

28	6	1.3	76	4	US-08-931-858E-22	Sequence 22, Appl
29	6	1.3	76	4	US-08-981-739-22	Sequence 22, Appl
30	6	1.3	83	3	US-08-673-814-4	Sequence 4, Appl1
31	6	1.3	83	4	US-09-115-824-4	Sequence 4, Appl1
32	6	1.3	90	4	US-09-103-478-29	Sequence 29, Appl
33	6	1.3	90	4	US-09-193-931C-29	Sequence 29, Appl
34	6	1.3	95	1	US-08-519-777-24	Sequence 24, Appl
35	6	1.3	95	1	US-08-742-035-24	Sequence 24, Appl
36	6	1.3	95	2	US-08-777-019-24	Sequence 24, Appl
37	6	1.3	95	2	US-08-777-143-24	Sequence 24, Appl
38	6	1.3	95	3	US-08-775-414-24	Sequence 24, Appl
39	6	1.3	95	4	US-08-931-858E-24	Sequence 24, Appl
40	6	1.3	95	4	US-08-981-739-24	Sequence 24, Appl
41	6	1.3	98	2	US-09-010-928B-20	Sequence 20, Appl
42	6	1.3	102	1	US-07-901-703-5	Sequence 5, Appl1
43	6	1.3	102	1	US-08-278-729A-11	Sequence 11, Appl
44	6	1.3	102	1	US-08-155-343A-11	Sequence 11, Appl
45	6	1.3	102	1	US-08-406-672-11	Sequence 11, Appl

ALIGNMENTS

```

RESULT 1
US-08-749-903-1
; Sequence 1, Application US/08749903
; Patent No. 5759812
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,903
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5759812e
; IMMEDIATE SOURCE:
; CLONE: 989953
; US-08-749-903-1

```

Query Match 100.0%; Score 472; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5759612e
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 374792
; US-08-749-903-3

Query Match      26.3%; Score 124; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.7e-112;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 LHTSHCLASEGVWISSLDYKNGKGGFVLLDDETFEVKGTWPRGGAAPLGIDFWYQPR 195
    |||
Db 136 LHTSHCLASEGVWISSLDYKNGKGGFVLLDDETFEVKGTWPRGGAAPLGIDFWYQPR 195
    |||
QY 196 HNWISTEMAAPVNLRDGFNPADVEAGLYGSHLYWMDMORHEIVQTLTKDGLIPLPIR 255
    |||
Db 196 HNWISTEMAAPVNLRDGFNPADVEAGLYGSHLYWMDMORHEIVQTLTKDGLIPLPIR 255
    |||
QY 256 LHNP 259
    |||
Db 256 LHNP 259

RESULT 4
US-09-088-641-3
; Sequence 3, Application US/09088641
; Patent No. 6312895
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,641
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/749,903
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

```

; TOPOLOGY: linear
; MOLECULE TYPE: No. 6312895e
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 374792
; US-09-088-641-3

Query Match      26.3%; Score 124; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.7e-112;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 LHTSHCLASEGVWISSLDYKNGKGGFVLLDDETFEVKGTWPRGGAAPLGIDFWYQPR 195
    |||
Db 136 LHTSHCLASEGVWISSLDYKNGKGGFVLLDDETFEVKGTWPRGGAAPLGIDFWYQPR 195
    |||
QY 196 HNWISTEMAAPVNLRDGFNPADVEAGLYGSHLYWMDMORHEIVQTLTKDGLIPLPIR 255
    |||
Db 196 HNWISTEMAAPVNLRDGFNPADVEAGLYGSHLYWMDMORHEIVQTLTKDGLIPLPIR 255
    |||
QY 256 LHNP 259
    |||
Db 256 LHNP 259

RESULT 5
US-08-749-903-4
; Sequence 4, Application US/08749903
; Patent No. 5759812
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,903
; FILING DATE: Filled Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 227630
; US-08-749-903-4

Query Match      10.4%; Score 49; DB 1; Length 472;
```

Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 CGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQY 56
DB 8 CGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQY 56

RESULT 6

US-09-088-641-4
; Sequence 4, Application US/09088641
; Patent No. 6312895
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,641
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/749,903
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PE-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 227630
US-09-088-641-4

Query Match 10.4%; Score 49; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.7e-39;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 CGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQY 56
DB 8 CGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQY 56

RESULT 7

US-08-749-903-5
; Sequence 5, Application US/08749903
; Patent No. 5759812
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US

ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,903
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PE-0163 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 298710
US-08-749-903-5

Query Match 9.3%; Score 44; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.2e-34;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 STPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQY 56
DB 13 STPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQY 56

RESULT 8

US-09-088-641-5
; Sequence 5, Application US/09088641
; Patent No. 6312895
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,641
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/749,903

FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Billings, Lucy J.
 REGISTRATION NUMBER: 36,749
 REFERENCE/DOCKET NUMBER: PF-0163 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555
 TELEFAX: 415-845-4166
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 472 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 IMMEDIATE SOURCE:
 LIBRARY: GenBank
 CLONE: 298710
 US-09-088-641-5

Query Match 9.3%; Score 44; DB 4; Length 472;
 Best Local Similarity 100.0%; Pred. No. 1.2e-34;
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 STELEAMKPREELIVLPCTIRNTGTEAPYLATVDVDPKSPQY 56
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 13 STELEAMKPREELIVLPCTIRNTGTEAPYLATVDVDPKSPQY 56

RESULT 9

US-08-858-207A-448
 ; Sequence 448, Application US/08858207A
 ; Patent No. 6348328

GENERAL INFORMATION:

APPLICANT: Black, Michael
 APPLICANT: Hodgson, John
 APPLICANT: Knowles, David
 APPLICANT: Nicholas, Richard
 APPLICANT: Stodola, Robert
 TITLE OF INVENTION: No. 6348328el Compounds
 NUMBER OF SEQUENCES: 552
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SmithKline Beecham Corporation
 STREET: 709 Swedeland Road
 CITY: King of Prussia
 STATE: PA
 COUNTRY: USA
 ZIP: 19406-0939

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/858,207A
 FILING DATE: 09-MAY-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/017670
 FILING DATE: 14-MAY-1996

ATTORNEY/AGENT INFORMATION:

NAME: Gimmil, Edward R.
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P50475
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090

TELEX:

INFORMATION FOR SEQ ID NO: 448:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 251 amino acids
 TYPE: amino acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: No. 6348328e
 US-08-858-207A-448

Query Match 1.5%; Score 7; DB 4; Length 251;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 260 DAAGFV 266
 |||||||
 Db 172 DAAGFV 178

RESULT 10

US-08-459-100A-3
 ; Sequence 3, Application US/08459100A
 ; Patent No. 5654172

GENERAL INFORMATION:

APPLICANT: Li, ET AL.
 TITLE OF INVENTION: GABA_A Receptor Epsilon Subunit
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
 ADDRESSEE: CECCHI, STEWART & OLSTEIN
 STREET: 6 BECKER FARM ROAD
 CITY: ROSELAND
 STATE: NEW JERSEY
 COUNTRY: USA
 ZIP: 07068

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 INCH DISKETTE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: MS-DOS
 SOFTWARE: WORD PERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/459,100A
 FILING DATE: June 2, 1995
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/09589
 FILING DATE: 26 AUG 1994

ATTORNEY/AGENT INFORMATION:

NAME: MULLINS, J.G.
 REGISTRATION NUMBER: 33,073
 REFERENCE/DOCKET NUMBER: 325800-325
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
 LENGTH: 467 AMINO ACIDS
 TYPE: AMINO ACID
 STRANDEDNESS:
 TOPOLOGY: LINEAR
 MOLECULE TYPE: PROTEIN
 US-08-459-100A-3

Query Match 1.5%; Score 7; DB 1; Length 467;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 PREELIV 28
 |||||||
 Db 214 PREELIV 220

RESULT 11

US-08-307-499-20
 ; Sequence 20, Application US/08307499
 ; Patent No. 5651972

```

; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/307,499
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-307-499-20

Query Match 1.5%; Score 7; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 TDILSL 317
DB 156 TDILSL 162

RESULT 12
US-09-299-268-20
; Sequence 20, Application US/09299268
; Patent No. 6217882
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Vi uela, Eladio
; APPLICANT: Gibbs, E.P.J.
; TITLE OF INVENTION: Use of Recombinant Swine Poxvirus as a
; TITLE OF INVENTION: Live Vaccine Vector
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David R. Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
```

```

; CITY: Gainesville
; STATE: Florida
; COUNTRY: U.S.A.
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/901,127
; FILING DATE:
; APPLICATION NUMBER: US 07/908,241
; FILING DATE: 1-JUL-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/908,630
; FILING DATE: 29-JUN-1992
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/342,212
; FILING DATE: 21-APR-1992
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 904-375-8100
; TELEFAX: 904-372-5800
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-299-268-20

Query Match 1.5%; Score 7; DB 4; Length 500;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 311 TDILSL 317
DB 156 TDILSL 162

RESULT 13
US-09-031-563-21
; Sequence 21, Application US/09031563A
; Patent No. 6022708
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage
; APPLICANT: Arnon Rosenthal
; TITLE OF INVENTION: Fused
; FILE REFERENCE: P1272
; CURRENT APPLICATION NUMBER: US/09/031,563A
; CURRENT FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 21
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-031-563-21

Query Match 1.5%; Score 7; DB 3; Length 685;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 450 PLGPALA 456
| | | | |
Db 555 PLGPALA 561

RESULT 14
US-08-907-166-6
; Sequence 6, Application US/08907166
; Patent No. 5948666
; GENERAL INFORMATION:
; APPLICANT: Callen, Walter
; APPLICANT: Mather, Eric
; TITLE OF INVENTION: ISOLATION AND IDENTIFICATION OF NOVEL POLYMERASES
; FILE REFERENCE: 09010/027001
; CURRENT APPLICATION NUMBER: US/08/907,166
; CURRENT FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 788
; TYPE: PRT
; ORGANISM: Archaeoglobus lithotrophicus
US-08-907-166-6

Query Match 1.5%; Score 7; DB 2; Length 788;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 297 KVKGWLL 303
| | | | |
Db 3 KVKGWLL 9

RESULT 15
US-09-031-563-27
; Sequence 27, Application US/09031563A
; Patent No. 6022708
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage
; APPLICANT: Arnon Rosenthal
; TITLE OF INVENTION: Fused
; FILE REFERENCE: P1272
; CURRENT APPLICATION NUMBER: US/09/031,563A
; CURRENT FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 27
; SEQ ID NO 27
; LENGTH: 1055
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-031-563-27

Query Match 1.5%; Score 7; DB 3; Length 1055;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 450 PLGPALA 456
| | | | |
Db 925 PLGPALA 931

Search completed: August 29, 2002, 16:52:25
Job time: 159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:51:06 ; Search time 108.86 Seconds
(Without alignments)
1526.133 Million cell updates/sec

Title: US-09-841-758-1
Perfect score: 472
Sequence: 1 MATKCGNGPGYSTPLFAMK PALAHLRYPGGDCSSDIWI 472

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 3502263 seqs, 351980561 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	472	100.0	472	22	US-09-841-758-1
2	472	100.0	472	22	US-09-981-353-11
3	472	100.0	472	26	US-60-239-841-11
4	457	96.8	493	1	PCT-US00-05988-11717
5	457	96.8	499	23	US-09-925-300-11717
6	335	71.0	514	26	US-60-340-187-420
7	124	26.3	472	22	US-09-841-758-3

8	102	21.6	582	1	PCT-US01-08656-6715
9	102	21.6	582	1	PCT-US01-14827-8378
10	77	16.3	138	1	PCT-US00-05989-879
11	77	16.3	138	23	US-09-925-297-879
12	58	12.3	135	20	US-09-621-976-6846
13	58	12.3	135	26	US-60-147-499-6846
14	44	10.4	472	22	US-09-841-758-4
15	44	9.3	472	22	US-09-841-758-5
16	16	3.4	199	20	US-09-620-394B-8286
17	16	3.4	236	20	US-09-620-394B-5717
18	16	3.4	236	20	US-09-620-394B-8285
19	16	3.4	275	20	US-09-620-394B-5716
20	16	3.4	275	20	US-09-620-394B-8284
21	16	3.4	328	20	US-09-620-394B-5715
22	16	3.4	409	21	US-09-708-427-21264
23	16	3.4	455	21	US-09-620-394B-6448
24	16	3.4	455	21	US-09-708-427-21263
25	16	3.4	480	26	US-60-324-109-20719
26	16	3.4	488	20	US-09-620-394B-6447
27	16	3.4	490	21	US-09-708-427-21262
28	16	3.4	503	20	US-09-620-394B-6446
29	15	3.2	58	1	PCT-US01-08656-6714
30	15	3.2	58	1	PCT-US01-14827-8377
31	15	3.2	400	21	US-09-708-427-11629
32	15	3.2	446	21	US-09-708-427-11628
33	15	3.2	455	20	US-09-620-394B-815
34	15	3.2	478	21	US-09-708-427-11627
35	15	3.2	487	20	US-09-620-394B-814
36	15	3.2	502	20	US-09-620-394B-813
37	14	3.0	142	21	US-09-708-427-38295
38	14	3.0	161	21	US-09-708-427-38294
39	14	3.0	493	26	US-60-324-109-20454
40	13	2.8	484	22	US-09-855-768-741
41	12	2.5	92	1	PCT-US01-08656-9360
42	10	2.1	323	19	US-09-595-329A-662
43	10	2.1	325	19	US-09-595-329A-661
44	10	2.1	339	19	US-09-595-329A-660
45	10	2.1	468	20	US-09-614-150-41685

ALIGNMENTS

RESULT 1
US-09-841-758-1
Sequence 1, Application US/09841758
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/841,758
FILING DATE: 24-Apr-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/088,641
FILING DATE: 1998-06-02
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0163 US

Sequence 6715, Ap
Sequence 8378, Ap
Sequence 879, App
Sequence 879, App
Sequence 6846, Ap
Sequence 6846, Ap
Sequence 4, Appl1
Sequence 5, Appl1
Sequence 8286, Ap
Sequence 5717, Ap
Sequence 8285, Ap
Sequence 8285, Ap
Sequence 5716, Ap
Sequence 8284, Ap
Sequence 5715, Ap
Sequence 21264, A
Sequence 6448, Ap
Sequence 21263, A
Sequence 20719, A
Sequence 6447, Ap
Sequence 21262, A
Sequence 6446, Ap
Sequence 6714, Ap
Sequence 8377, Ap
Sequence 11629, A
Sequence 11628, A
Sequence 815, App
Sequence 11627, A
Sequence 814, App
Sequence 813, App
Sequence 38295, A
Sequence 38294, A
Sequence 20454, A
Sequence 741, App
Sequence 9360, App
Sequence 662, App
Sequence 661, App
Sequence 660, App
Sequence 41685, A

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 472 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: None
;   IMMEDIATE SOURCE:
;   LIBRARY: <Unknown>
;   CLONE: 989953
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-841-758-1

Query Match      100.0%; Score 472; DB 22; Length 472;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCITYRMTGEAPDYLATVDVDPKSPQYCVI 60
DB 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCITYRMTGEAPDYLATVDVDPKSPQYCVI 60
QY 61 HRLPMPNLKDELHHSQWNTSCSCFGDSTKSRKLVLPSLISSRIYVVDVGSPPRAPKLHK 120
DB 61 HRLPMPNLKDELHHSQWNTSCSCFGDSTKSRKLVLPSLISSRIYVVDVGSPPRAPKLHK 120
QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
DB 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
DB 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
QY 181 GGAAPLGYDFWYQPRHNWVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMQRIEIVQ 240
DB 181 GGAAPLGYDFWYQPRHNWVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMQRIEIVQ 240
QY 181 GGAAPLGYDFWYQPRHNWVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMQRIEIVQ 240
DB 181 GGAAPLGYDFWYQPRHNWVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMQRIEIVQ 240
QY 241 TSLKDGILPLEIRLHNPDAAGFVGALSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
DB 241 TSLKDGILPLEIRLHNPDAAGFVGALSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
QY 301 WLLPMPGLITDILSLDRFLYFSNMHLHGLRDYDISDPQRRLTGOLFGLGSIVKGGP 360
DB 301 WLLPMPGLITDILSLDRFLYFSNMHLHGLRDYDISDPQRRLTGOLFGLGSIVKGGP 360
QY 361 VOVLDEELKSQPEPLVYVKGKRVAGPQMIQLSLDGKRLYITTSLSYAMDQKOFYDLIRE 420
DB 361 VOVLDEELKSQPEPLVYVKGKRVAGPQMIQLSLDGKRLYITTSLSYAMDQKOFYDLIRE 420
QY 421 GSVMLQVDVDTYKGLKLPNPLVDFGKEPLGPAALAHLELRYPGGCSDIWI 472
DB 421 GSVMLQVDVDTYKGLKLPNPLVDFGKEPLGPAALAHLELRYPGGCSDIWI 472

```

```

RESULT 2
US-09-981-353-11
; Sequence 11, Application US/09981353
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO: 11
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 611514CD1

```

```

US-09-981-353-11

Query Match      100.0%; Score 472; DB 23; Length 472;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCITYRMTGEAPDYLATVDVDPKSPQYCVI 60
DB 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCITYRMTGEAPDYLATVDVDPKSPQYCVI 60
QY 61 HRLPMPNLKDELHHSQWNTSCSCFGDSTKSRKLVLPSLISSRIYVVDVGSPPRAPKLHK 120
DB 61 HRLPMPNLKDELHHSQWNTSCSCFGDSTKSRKLVLPSLISSRIYVVDVGSPPRAPKLHK 120
QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
DB 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
DB 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLDGKNGKGGFVLLDGETFEVKTWERP 180
QY 181 GGAAPLGYDFWYQPRHNWVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMQRIEIVQ 240
DB 181 GGAAPLGYDFWYQPRHNWVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMQRIEIVQ 240
QY 241 TSLKDGILPLEIRLHNPDAAGFVGALSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
DB 241 TSLKDGILPLEIRLHNPDAAGFVGALSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
QY 301 WLLPMPGLITDILSLDRFLYFSNMHLHGLRDYDISDPQRRLTGOLFGLGSIVKGGP 360
DB 301 WLLPMPGLITDILSLDRFLYFSNMHLHGLRDYDISDPQRRLTGOLFGLGSIVKGGP 360
QY 361 VOVLDEELKSQPEPLVYVKGKRVAGPQMIQLSLDGKRLYITTSLSYAMDQKOFYDLIRE 420
DB 361 VOVLDEELKSQPEPLVYVKGKRVAGPQMIQLSLDGKRLYITTSLSYAMDQKOFYDLIRE 420
QY 421 GSVMLQVDVDTYKGLKLPNPLVDFGKEPLGPAALAHLELRYPGGCSDIWI 472
DB 421 GSVMLQVDVDTYKGLKLPNPLVDFGKEPLGPAALAHLELRYPGGCSDIWI 472

```

```

RESULT 3
US-60-239-841-11
; Sequence 11, Application US/60239841
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 P
; CURRENT APPLICATION NUMBER: US/60/239.841
; CURRENT FILING DATE: 2000-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO: 11
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 611514CD1
; US-60-239-841-11

```

```

Query Match      100.0%; Score 472; DB 26; Length 472;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCITYRMTGEAPDYLATVDVDPKSPQYCVI 60
DB 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCITYRMTGEAPDYLATVDVDPKSPQYCVI 60
QY 61 HRLPMPNLKDELHHSQWNTSCSCFGDSTKSRKLVLPSLISSRIYVVDVGSPPRAPKLHK 120
DB 61 HRLPMPNLKDELHHSQWNTSCSCFGDSTKSRKLVLPSLISSRIYVVDVGSPPRAPKLHK 120

```

QY 1 MAAGGNGGPGYSPPLPAMKGPREEIYYLPCTYNNCTEADPYLATADVPKSPQYCYI 60

Db 28 MATKGGNGGPGYSPPLPAMKGPREEIYYLPCTYNNCTEADPYLATADVPKSPQYCYI 87

QY 61 HRLPMPNLKDELHSHGWNCCSCFQDSTKSTKLVLPSSLISRIYVDVSGSEPRAPKLHK 120

Db 88 HRLPMPNLKDELHSHGWNCCSCFQDSTKSTKLVLPSSLISRIYVDVSGSEPRAPKLHK 147

QY 121 VIEPKDIHAKELAFHTSHCLAGFVWISSLDGVKNGGNGGVLDGEFFEVKGTWERP 180

[illegible]

```

Db 208 GGAAPGDFWQPRHNWVISTEMAAPNVLKRGFNADVEAGLSHYLVWDMORHEITQ 267
QY 241 TSLKDGILPIELIRLHNDAQGFVGCALSTIOREYKNEGTSVEKVIQVPPKVK 300
Db 268 TSLKDGILPIELIRLHNDAQGFVGCALSTIOREYKNEGTSVEKVIQVPPKVK 327
QY 301 WLPEMGLITDILSLDRFLYFSNMWHLGDLROYDISDPQRRLTGQLFGSIYKGP 360
Db 328 WLPEMGLITDILSLDRFLYFSNMWHLGDLROYDISDPQRRLTGQLFGSIYKGP 387
QY 361 VVLEDEELKSQPEPLVYKGRVAGPQMIOQLSDGKRLYITTSLSXANDKQFYDLIRE 420
Db 388 VVLEDEELKSQPEPLVYKGRVAGPQMIOQLSDGKRLYITTSLSXANDKQFYDLIRE 447
QY 421 GSVMLQVDVDTYKGGKLNPNFLVDFGKEPLGALAH 457
Db 448 GSVMLQVDVDTYKGGKLNPNFLVDFGKEPLGALAH 484

```

RESULT 6
US-60-340-187-420

Sequence 420, Application US/60340187

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Wang, Zhiwei

APPLICANT: Wang, Jian-Rui

APPLICANT: Ghosh, Malabika

APPLICANT: Meng, Gezh

APPLICANT: Boyle, Bryan J

APPLICANT: Drmanac, Radoje T

TITLE OF INVENTION: Novel Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT FILING DATE: 2001-12-12

CURRENT FILING DATE: 2001-12-12

PRIOR FILING DATE: 2000-01-21

PRIOR FILING DATE: 2000-01-21

PRIOR FILING DATE: 2000-04-25

PRIOR FILING DATE: 2000-04-25

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-12-22

PRIOR FILING DATE: 2000-01-25

PRIOR FILING DATE: 2000-01-25

PRIOR FILING DATE: 2001-01-25

PRIOR FILING DATE: 2001-01-25

PRIOR FILING DATE: 2000-02-03

PRIOR FILING DATE: 2000-02-03

PRIOR FILING DATE: 2000-04-27

PRIOR FILING DATE: 2000-04-27

PRIOR FILING DATE: 2001-02-05

PRIOR FILING DATE: 2001-02-05

PRIOR FILING DATE: 2000-02-28

PRIOR FILING DATE: 2000-02-28

PRIOR FILING DATE: 2000-05-18

PRIOR FILING DATE: 2000-05-18

PRIOR FILING DATE: 2001-02-26

PRIOR FILING DATE: 2001-02-26

PRIOR FILING DATE: 2000-03-07

PRIOR FILING DATE: 2000-03-07

PRIOR FILING DATE: 2000-05-19

PRIOR FILING DATE: 2000-05-19

PRIOR FILING DATE: 2001-03-05

PRIOR FILING DATE: 2001-03-05

PRIOR FILING DATE: 2000-03-31

PRIOR FILING DATE: 2000-03-31

PRIOR FILING DATE: 2000-08-23

PRIOR FILING DATE: 2000-08-23

PRIOR FILING DATE: 2001-03-30

PRIOR FILING DATE: 2001-03-30

PRIOR FILING DATE: 2000-04-18

PRIOR FILING DATE: 2000-04-18

PRIOR APPLICATION NUMBER: US 09/770,160

```

: PRIOR FILING DATE: 2001-01-26
: PRIOR APPLICATION NUMBER: PCT/US01/08656
: PRIOR FILING DATE: 2001-04-18
: PRIOR APPLICATION NUMBER: US 09/577,408
: PRIOR FILING DATE: 2000-05-18
: PRIOR APPLICATION NUMBER: PCT/US01/14827
: PRIOR FILING DATE: 2001-05-16
: NUMBER OF SEQ ID NOS: 1192
: SOFTWARE: pc-fl_genes Version 6.0
: SEQ ID NO 420
: LENGTH: 514
: TYPE: PRT
: ORGANISM: Homo sapiens
US-60-340-187-420

```

Query Match 71.0%; Score 335; DB 26; Length 514;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 75 GSWNTSSCFDSTKSRKLVPLSLISRIYVVDGSEPRAPKLHVKIDIAKCELA 134
Db 117 GSWNTSSCFDSTKSRKLVPLSLISRIYVVDGSEPRAPKLHVKIDIAKCELA 176
QY 135 FLHTSHCLASGEVMTISLGDVGNKGGFVLLDGETFEYKGTWERPGGAAPLGDFWQ 194
Db 177 FLHTSHCLASGEVMTISLGDVGNKGGFVLLDGETFEYKGTWERPGGAAPLGDFWQ 236
QY 195 RHNVISTEMAAPNVLKRGFNADVEAGLYGSHLVWDMORHEIVQTSLSKDGILPIELR 254
Db 237 RHNVISTEMAAPNVLKRGFNADVEAGLYGSHLVWDMORHEIVQTSLSKDGILPIELR 296
QY 255 FLHNDAAQGFVGCALSTIOREYKNEGTSVEKVIQVPPKVKGMILLPEMGLITDIL 314
Db 297 FLHNDAAQGFVGCALSTIOREYKNEGTSVEKVIQVPPKVKGMILLPEMGLITDIL 356
QY 315 LSLDRFLYFSNMWHLGDLROYDISDPQRRLTGQLFGSIYKGPVVOVLEDEELKSQPE 374
Db 357 LSLDRFLYFSNMWHLGDLROYDISDPQRRLTGQLFGSIYKGPVVOVLEDEELKSQPE 416
QY 375 PLVYKGRVAGPQMIOQLSDGKRLYITTSLSXAN 409
Db 417 PLVYKGRVAGPQMIOQLSDGKRLYITTSLSXAN 451

```

RESULT 7
US-09-841-758-3

Sequence 3, Application US/09841758

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESS: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/841,758

FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/088,641

FILING DATE: 1998-06-02

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0163 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: None
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1374792
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-841-758-3

Query Match 26.3%; Score 124; DB 22; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.1e-116;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 LHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGTERPGGAAPLGYDFWYQPR 195
|||||
DB 136 LHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGTERPGGAAPLGYDFWYQPR 195
QY 196 HNVVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMORHEIVOTLSLKDGLPLEIRF 255
|||||
DB 196 HNVVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMORHEIVOTLSLKDGLPLEIRF 255
QY 256 LHP 259
|||||
DB 256 LHP 259
RESULT 8
PCT-US01-08656-6715
; Sequence 6715, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 6715
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(582)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08656-6715

Query Match 21.6%; Score 102; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4e-94;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AFLTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGTERPGGAAPLGYDFWYQ 193
|||||
DB 175 AFLTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGTERPGGAAPLGYDFWYQ 234
QY 194 PRHNVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMOR 235
|||||
DB 235 PRHNVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMOR 276

RESULT 9
PCT-US01-14827-8378
; Sequence 8378, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 8378
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(582)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-14827-8378

Query Match 21.6%; Score 102; DB 1; Length 582;
Best Local Similarity 100.0%; Pred. No. 3.4e-94;
Matches 102; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 134 AFLTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGTERPGGAAPLGYDFWYQ 193
|||||
DB 175 AFLTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGTERPGGAAPLGYDFWYQ 234
QY 194 PRHNVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMOR 235
|||||
DB 235 PRHNVISTEMAAPNVLRDGFNPADVEAGLYGSHLYWDMOR 276

RESULT 10
PCT-US00-05989-879
; Sequence 879, Application PC/TUS0005989
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Pancreas and Pancreatic Cancer Associated Gene Sequences
; FILE REFERENCE: PA105PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05989
; CURRENT FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124,270
; EARLIER FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 879
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (102)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (111)

```
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (113)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (115)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (112)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (132)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: SITE
LOCATION: (135)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05989-879
```

```
Query Match 16.3%; Score 77; DB 1; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 132 ELAFLHTSHCLASGEVMTISLGDVNGKNGKGFVLLDGETFEVKGTWERPGGAAPLGDFW 191
|||
Db 6 ELAFLHTSHCLASGEVMTISLGDVNGKNGKGFVLLDGETFEVKGTWERPGGAAPLGDFW 65
```

```
Oy 192 YOPRHNMISTEWAAPN 208
|||
Db 66 YOPRHNMISTEWAAPN 82
```

```
RESULT 11
US-09-925-297-879
; Sequence 879, Application US/09925297
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA105
; CURRENT APPLICATION NUMBER: US/09/925,297
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 879
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (83)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (102)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (107)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
```

```
NAME/KEY: SITE
LOCATION: (113)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (115)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (125)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (132)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (135)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-879
```

```
Query Match 16.3%; Score 77; DB 23; Length 138;
Best Local Similarity 100.0%; Pred. No. 2.7e-69;
Matches 77; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 132 ELAFLHTSHCLASGEVMTISLGDVNGKNGKGFVLLDGETFEVKGTWERPGGAAPLGDFW 191
|||
Db 6 ELAFLHTSHCLASGEVMTISLGDVNGKNGKGFVLLDGETFEVKGTWERPGGAAPLGDFW 65
```

```
Oy 192 YOPRHNMISTEWAAPN 208
|||
Db 66 YOPRHNMISTEWAAPN 82
```

```
RESULT 12
US-09-621-976-6846
; Sequence 6846, Application US/09621976
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 6846
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6846
```

```
Query Match 12.3%; Score 58; DB 20; Length 135;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 MATKCGNGPGSTPLEAMKGPREEIVYLPCYRNTGTGAPDYLATVVDVDRKSPQYQ 58
|||
Db 1 MATKCGNGPGSTPLEAMKGPREEIVYLPCYRNTGTGAPDYLATVVDVDRKSPQYQ 58
```

```
RESULT 13
US-60-147-499-6846
; Sequence 6846, Application US/60147499
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTS and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/60/147,499
```

;; CURRENT FILING DATE: 1999-08-05
;; NUMBER OF SEQ ID NOS: 19335
;; SOFTWARE: Patent.pm
;; SEQ ID NO 6846
;; LENGTH: 135
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-60-147-499-6846

Query Match 12.3%; Score 58; DB 26; Length 135;
Best Local Similarity 100.0%; Pred. No. 5.9e-50;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MATCGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGTGTEADPYLATVDVDPKSPQYCO 58
DB 1 MATCGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGTGTEADPYLATVDVDPKSPQYCO 58

RESULT 14
US-09-841-758-4
; Sequence 4, Application us/09841758
; GENERAL INFORMATION:

APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: us/09/841,758
FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/088,641
FILING DATE: 1998-06-02
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0163 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 472 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 227630

SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-841-758-4

Query Match 10.4%; Score 49; DB 22; Length 472;
Best Local Similarity 100.0%; Pred. No. 2.7e-40;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 CGGCGYSTPLEAMKGPREEIYVLPCTYRNTGTGTEADPYLATVDVDPKSPQY 56
DB 8 CGGCGYSTPLEAMKGPREEIYVLPCTYRNTGTGTEADPYLATVDVDPKSPQY 56

RESULT 15
US-09-841-758-5
; Sequence 5, Application us/09841758
; GENERAL INFORMATION:

APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: us/09/841,758

FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/088,641

FILING DATE: 1998-06-02

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0163 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 472 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 298710

SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-841-758-5

Query Match 9.3%; Score 44; DB 22; Length 472;
Best Local Similarity 100.0%; Pred. No. 3.4e-35;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 STPLEAMKGPREEIYVLPCTYRNTGTGTEADPYLATVDVDPKSPQY 56
DB 13 STPLEAMKGPREEIYVLPCTYRNTGTGTEADPYLATVDVDPKSPQY 56

Search completed: August 29, 2002, 16:54:47
Job time: 221 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:51:26 ; Search time 37 Seconds
(without alignments)
3094.797 Million cell updates/sec

Title: US-09-841-758-1
472

Sequence: 1 MATKCGNCGPGYSTPLEAMK.....PALAHLRYPGGDCSSDIWI 472

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 751360 seqs, 242600617 residues

Word size : 0

Total number of hits satisfying chosen parameters: 751360

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Pending_Patents_AA_New:*
1: /cgn2_6/ptodata/2/paa/pcr_NEW.COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_NEW.COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_NEW.COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_NEW.COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US09_NEW.COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US10_NEW.COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US60_NEW.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	217	46.0	235	US-09-629-469A-18237	Sequence 18237, A
2	16	3.4	167	US-09-935-625-9670	Sequence 9670, Ap
3	16	3.4	236	US-09-935-625-10394	Sequence 10394, A
4	16	3.4	236	US-09-935-625-11323	Sequence 11323, A
5	16	3.4	275	US-09-935-625-10383	Sequence 10383, A
6	16	3.4	275	US-09-935-625-11322	Sequence 11322, A
7	15	3.2	200	US-09-935-625-5811	Sequence 5811, Ap
8	15	3.2	200	US-09-935-625-10423	Sequence 10423, A
9	15	3.2	239	US-09-935-625-5810	Sequence 5810, Ap
10	15	3.2	239	US-09-935-625-10422	Sequence 10422, A
11	15	3.2	291	US-09-935-625-5809	Sequence 5809, Ap
12	15	3.2	291	US-09-935-625-30421	Sequence 30421, A
13	15	3.2	400	US-09-935-625-5071	Sequence 5071, Ap
14	15	3.2	409	US-09-935-625-4908	Sequence 4908, Ap
15	15	3.2	412	US-09-935-625-5884	Sequence 5884, Ap
16	15	3.2	412	US-09-935-625-10791	Sequence 10791, A
17	15	3.2	446	US-09-935-625-5070	Sequence 5070, Ap
18	15	3.2	458	US-09-935-625-5883	Sequence 5883, Ap
19	15	3.2	458	US-09-935-625-30790	Sequence 30790, A
20	15	3.2	469	US-09-935-625-4907	Sequence 4907, A
21	15	3.2	478	US-09-935-625-5069	Sequence 5069, Ap
22	15	3.2	490	US-09-935-625-5882	Sequence 5882, Ap
23	15	3.2	490	US-09-935-625-30789	Sequence 30789, A
24	15	3.2	502	US-09-935-625-4906	Sequence 4906, Ap
25	14	3.0	492	US-60-391-781-1043	Sequence 1043, Ap
26	11	2.3	219	US-09-935-625-10395	Sequence 10395, A

27	11	2.3	219	US-09-935-625-31324	Sequence 31324, A
28	10	2.1	163	US-09-935-625-9671	Sequence 9671, Ap
29	10	2.1	455	US-09-935-625-2952	Sequence 2952, Ap
30	10	2.1	455	US-09-935-625-12244	Sequence 12244, A
31	10	2.1	455	US-09-935-625-16885	Sequence 16885, A
32	10	2.1	455	US-09-935-625-28653	Sequence 28653, A
33	10	2.1	476	US-09-935-625-12243	Sequence 12243, A
34	10	2.1	476	US-09-935-625-16884	Sequence 16884, A
35	10	2.1	476	US-09-935-625-28652	Sequence 28652, A
36	10	2.1	480	US-09-935-625-2950	Sequence 2950, Ap
37	10	2.1	480	US-09-935-625-12242	Sequence 12242, A
38	10	2.1	480	US-09-935-625-16883	Sequence 16883, A
39	10	2.1	480	US-09-935-625-28651	Sequence 28651, A
40	10	2.1	487	US-09-791-537-35612	Sequence 35612, A
41	10	2.1	487	US-60-360-039-6062	Sequence 6062, Ap
42	10	2.1	487	US-09-791-537-3940	Sequence 3940, Ap
43	9	1.9	286	US-09-791-537-740	Sequence 740, Ap
44	7	1.5	86	US-10-144-702-828	Sequence 828, Ap
45	7	1.5	86	PCT-US02-09785-828	

ALIGNMENTS

Result 1
US-09-629-469A-18237
Sequence 18237, Application US/09629469A
GENERAL INFORMATION:
APPLICANT: OTA, TOSHIO
APPLICANT: ISOGAI, TAKAO
APPLICANT: NISHIKAWA, TETSUO
APPLICANT: HAYASHI, KOJI
APPLICANT: SAITO, KAORI
APPLICANT: YAMAMOTO, JUNICHI
APPLICANT: ISHII, SHIZUKO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: WAKAMATSU, AI
APPLICANT: NAGAI, KEIICHI
APPLICANT: OTSUKI, TETSUJI
TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH cDNA AND THEIR USE
FILE REFERENCE: 084335/0123
CURRENT APPLICATION NUMBER: US/09/629, 469A
CURRENT FILING DATE: 2000-07-28
PRIOR APPLICATION NUMBER: JP 1999-248036
PRIOR FILING DATE: 1999-07-29
PRIOR APPLICATION NUMBER: JP 1999-300253
PRIOR FILING DATE: 1999-08-27
PRIOR APPLICATION NUMBER: JP 2000-118776
PRIOR FILING DATE: 2000-01-11
PRIOR APPLICATION NUMBER: JP 2000-183767
PRIOR FILING DATE: 2000-05-02
PRIOR APPLICATION NUMBER: JP 2000-241899
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/159, 590
PRIOR FILING DATE: 1999-10-18
PRIOR APPLICATION NUMBER: 60/183, 322
PRIOR FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 19025
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 18237
LENGTH: 235
TYPE: PRT
ORGANISM: Homo sapiens
US-09-629-469A-18237

Query Match 46.0%; Score 217; DB 5; Length 235;
Best Local Similarity 100.0%; Pred. No. 5.8e-211; Indels 0; Gaps 0;
Matches 217; Conservative 0; Mismatches 0;

65 MFNLDLHSHGWNCCSCFDSRTKRLVLPGLSSRLVYDVGSEPRAPRLHVIIEP 124
|||||
1 MFNLDLHSHGWNCCSCFDSRTKRLVLPGLSSRLVYDVGSEPRAPRLHVIIEP 60
|||||

Fri Aug 30 09:21:30 2002

us-09-841-758-1.rapn

Page 4

```
RESULT 11
US-09-935-625-5809
; Sequence 5809, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 5809
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..291
; OTHER INFORMATION: Ceres Seq. ID no. 3072810
US-09-935-625-5809
```

```
Query Match          3.2%; Score 15; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 386 GPOMIQLSLDGKRLY 400
    |||||||
Db 205 GPOMIQLSLDGKRLY 219
```

```
RESULT 12
US-09-935-625-30421
; Sequence 30421, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 30421
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..291
; OTHER INFORMATION: Ceres Seq. ID no. 3072810
US-09-935-625-30421
```

```
Query Match          3.2%; Score 15; DB 5; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 386 GPOMIQLSLDGKRLY 400
    |||||||
Db 205 GPOMIQLSLDGKRLY 219
```

```
RESULT 13
US-09-935-625-5071
; Sequence 5071, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
```

```
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 5071
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..400
; OTHER INFORMATION: Ceres Seq. ID no. 1823506
US-09-935-625-5071
```

```
Query Match          3.2%; Score 15; DB 5; Length 400;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 386 GPOMIQLSLDGKRLY 400
    |||||||
Db 314 GPOMIQLSLDGKRLY 328
```

```
RESULT 14
US-09-935-625-4908
; Sequence 4908, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 4908
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..409
; OTHER INFORMATION: Ceres Seq. ID no. 1441047
US-09-935-625-4908
```

```
Query Match          3.2%; Score 15; DB 5; Length 409;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 386 GPOMIQLSLDGKRLY 400
    |||||||
Db 323 GPOMIQLSLDGKRLY 337
```

```
RESULT 15
US-09-935-625-5884
; Sequence 5884, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 5884
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..412
; OTHER INFORMATION: Ceres Seq. ID no. 3089925
US-09-935-625-5884
```

Query Match 3.2%; Score 15; DB 5; Length 412;
Best Local Similarity 100.0%; Pred. No. 2.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 GPQMQLSLDGRLY 400
|||||
Db 326 GPQMQLSLDGRLY 340

Search completed: August 29, 2002, 16:55:31
Job time: 245 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:50:26 ; Search time 20.94 Seconds
(without alignments)
2165.910 Million cell updates/sec

Title: US-09-841-758-1
472

Sequence: 1 MATKCGNCGPSTPLEAMK.....PALAHLRLYPGGDCSSDIWI 472

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

PIR 71:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	26.3	472	2	G01872
2	49	10.4	472	2	S27878
3	16	3.4	490	2	D71401
4	15	3.2	478	2	A85153
5	15	3.2	1070	2	E71401
6	10	2.1	471	2	T32635
7	10	2.1	487	1	T32635
8	9	1.9	286	1	RLPUGG
9	8	1.7	85	2	S56097
10	8	1.7	148	2	H69813
11	8	1.7	240	2	H70248
12	8	1.7	240	2	F70250
13	8	1.7	345	2	T27695
14	8	1.7	429	2	D95911
15	8	1.7	442	2	T50687
16	8	1.7	504	2	S74034
17	8	1.7	725	2	AH1428
18	7	1.5	72	2	T18606
19	7	1.5	72	2	T29015
20	7	1.5	77	2	H71206
21	7	1.5	132	2	G75466
22	7	1.5	133	2	S63211
23	7	1.5	168	2	S44782
24	7	1.5	179	2	A64846
25	7	1.5	179	2	E90805
26	7	1.5	179	2	B85655
27	7	1.5	201	2	D70556
28	7	1.5	207	2	D90223
29	7	1.5	207	2	T26455

30	7	1.5	212	2	A44994	eggshe1 protein 1
31	7	1.5	227	2	T26623	hypothetical prote
32	7	1.5	228	2	S76876	hypothetical prote
33	7	1.5	255	2	C75392	oxidoreductase, sh
34	7	1.5	264	1	MMBV1.1	30K protein - toma
35	7	1.5	266	1	S55397	probable methyltra
36	7	1.5	278	2	G82074	mazg protein VC245
37	7	1.5	286	2	A95147	DNA processing pro
38	7	1.5	286	2	G98014	DNA processing Smf
39	7	1.5	305	2	F84595	hypothetical prote
40	7	1.5	305	2	H97323	TPR-repeat-contain
41	7	1.5	308	2	H70974	probable nucleosid
42	7	1.5	316	2	C75205	hypothetical prote
43	7	1.5	325	2	S18575	sym protein - Rhl
44	7	1.5	326	2	A44505	sym protein - Rhl
45	7	1.5	326	2	F95319	sym transcription

ALIGNMENTS

RESULT 1
G01872
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Jun-2000
C/Accession: G01872
R.Chang, P.W.G.
Submitted to the EMBL Data Library, June 1995
A/Reference number: H00690
A/Accession: G01872
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A:Residues: 1-472 <HAN>
A/Cross-references: EMBL:029091; NID:q1374791; PIDN:AAB02395.1; PID:q1374792
C/Genetics:
A:Gene: hsbP
C/Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 26.3%; Score 124; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 66-117;
Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 LHTSHCLASGEVMTSSIGDYKNGKNGGFEVLLDGETFEYKGTWEPGGAAPLGDFWQPR 195
Db 136 LHTSHCLASGEVMTSSIGDYKNGKNGGFEVLLDGETFEYKGTWEPGGAAPLGDFWQPR 195

QY 196 HNWISTEMAPNVLRDGFNPADVEAGLYGSHLYVMDQRHEIVQTLSLKDGILPIEIRF 255
Db 196 HNWISTEMAPNVLRDGFNPADVEAGLYGSHLYVMDQRHEIVQTLSLKDGILPIEIRF 255

QY 256 LHPN 259
Db 256 LHPN 259

RESULT 2
S27878
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Jun-2000
C/Accession: S27878; JN0271
R.Bansal, M.P.; Cook, R.; Oborn, C.J.; Scott, J.; Mukhopadhyay, T.; Medina, D.
Submitted to the EMBL Data Library, February 1991
A/Description: Molecular cloning and sequencing of a liver protein which binds seleni
A/Reference number: S27878
A/Accession: S27878
A/Molecule type: mRNA
A:Residues: 1-472 <BAN>
A/Cross-references: EMBL:M32032; NID:g200951; PIDN:AAA40104.1; PID:g200952
R.Pumford, N.R.; Martin, B.M.; Hinson, J.A.
Biochem. Biophys. Res. Commun. 182, 1348-1355, 1992

A:Title: A metabolite of acetaminophen covalently binds to the 56 kDa selenium binding H
A:Reference number: JN0271; MUID:92171951
A:Accession: JN0271
A:Molecule type: protein
A:Residues: 72-73, 'K', 175-181, 'G', 183, 'S', 185-189, 196-220, 228-242, 290-296, 334-343, 393-40
C:Comment: This protein covalently binds a metabolite of acetaminophen. This binding co
C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5
C:Keywords: liver; selenium binding

Query Match 10.4%; Score 49; DB 2; Length 472;
Best Local Similarity 100.0%; Pred. No. 4,7e-41;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 CGPGSTPLEAMKGPREEVYLPCTYRMTGTPAPYLATVDVDPKSPQY 56
|||||
DB 8 CGPGSTPLEAMKGPREEVYLPCTYRMTGTPAPYLATVDVDPKSPQY 56

RESULT 3
D71401

probable selenium-binding protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000

C:Accession: D71401

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Woltzenegger, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; An
C.; Chalwatizis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113

A:Accession: D71401

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-490 <BEV>

A:Cross-references: GB:297335; NID:g2244747; PID:g2244759

C:Genetics:

A:Map position: 4COP9-4G3845

C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 3.4%; Score 16; DB 2; Length 490;
Best Local Similarity 100.0%; Pred. No. 1,2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 385 GGPOMIQLSLDGKRLY 400
|||||
DB 403 GGPOMIQLSLDGKRLY 418

RESULT 4
A85153

selenium-binding protein like [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 02-Mar-2001

C:Accession: A85153

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488

A:Accession: A85153

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-478 <STO>

A:Cross-references: GB:NC_001268; NID:g7268108; PIDN:CAB78446.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4G14040

A:Map position: 4

C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 3.2%; Score 15; DB 2; Length 478;
Best Local Similarity 100.0%; Pred. No. 1,2e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 GPMIQLSLDGKRLY 400
|||||
DB 392 GPMIQLSLDGKRLY 406

RESULT 5
E71401

probable selenium-binding protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variety: Columbia

C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 11-Jan-2002

C:Accession: E71401

R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Woltzenegger, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalwatizis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113

A:Accession: E71401

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1070 <BEV>

A:Cross-references: GB:297335; NID:g2244747; PID:g2244760

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match 3.2%; Score 15; DB 2; Length 1070;
Best Local Similarity 100.0%; Pred. No. 2,4e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 GPMIQLSLDGKRLY 400
|||||
DB 984 GPMIQLSLDGKRLY 998

RESULT 6
T26612

hypothetical protein Y37A1B.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

C:Accession: T26612

R:McMurray, A.
submitted to the EMBL Data Library, June 1998

A:Reference number: Z20245

A:Accession: T26612

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-471 <WIL>

A:Cross-references: EMBL:AL023835; PIDN:CAA19490.1; GSPDB:GN00022; CESP:Y37A1B.5

A:Experimental source: clone Y37A1B

C:Genetics:

A:Gene: CESP:Y37A1B.5

A:Map position: 4

A:introns: 58/3; 91/2; 215/1; 282/3; 418/2

C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 2.1%; Score 10; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 0,14;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 391 QLSLDGKRLY 400
|||||

Db 390 QLSLDGKRLY 399

RESULT 7

T32635 phosphoprotein phosphatase (EC 3.1.3.16) F42G8.8 [similarity] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Nov-2001

C/Accession: T32635

R/Gating, S.; Holmes, A. submitted to the EMBL data library, December 1997

A/Description: The sequence of C. elegans cosmid F42G8.

A/Reference number: Z21203

A/Accession: T32635

A/Status: translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-487 <GAT>

A/Cross-references: EMBL:AF036618, PTDN:MA92072.1; GSPDB:GN00022; CESP:F42G8.8

C/Experimental source: strain Bristol N2; clone F42G8

C/Comment: This sequence shares a domain (approximately residues 378-487) with mammalian C/gene: CESP:F42G8.8

A/Map position: 4

A/Tritons: 27/3; 97/2; 122/3; 203/3; 236/2; 274/3; 302/3; 377/1

C/Superfamily: Caenorhabditis elegans phosphoprotein phosphatase F42G8.8; phosphoesterase C/Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific F/30-314/Domain: phosphoprotein phosphatase homology <PPP>

F/58-137/Domain: phosphoesterase core homology <PPEC>

F/64,66,103/Binding site: iron (Asp, His, Asp) #status predicted

F/103,135,188,273/Binding site: zinc (Asp, Asn, His, His) #status predicted

F/106,136,297/Active site: Asp, His, Tyr #status predicted

F/107,246/Binding site: substrate phosphate (Arg) #status predicted

Query Match 2.1%; Score 10; DB 1; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.14; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 391 QLSLDGKRLY 400
|||||
Db 459 QLSLDGKRLY 468

RESULT 8

RLPUGG RNA N-glycosidase (EC 3.2.2.22) alpha-momorcharin precursor [validated] - balsam pear

N/Alternate names: agglutinin, momordin A; ribosome-inactivating protein momorcharin alp

C/Species: Momordica charantia (balsam pear, bitter melon)

C/Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 15-Sep-2000

C/Accession: S14273; A61318; S16490; JN0628; S01670

R/Ho, W.K.K.; Liu, S.C.; Shaw, P.C.; Yeung, H.W.; Ng, T.B.; Chan, W.Y.

Biochim. Biophys. Acta 1088, 311-314, 1991

A/Title: Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating protein.

A/Reference number: S14273; MUID:91159486

A/Accession: S14273

A/Molecule type: mRNA

A/Residues: 1-286 <HOW>

A/Cross-references: EMBL:X57682; NID:q19527; PTDN:CAA40869.1; PID:q19528

R/Li, S.S.L.

Experientia 36, 524-527, 1980

A/Title: Purification and partial characterization of two lectins from Momordica charantia

A/Reference number: A61318; MUID:80201763

A/Accession: A61318

A/Molecule type: protein

A/Residues: 24-68, 'X', '70' <MON>

R/Minami, Y.; Funatsu, G.

Biosci. Biotechnol. Biochem. 57, 1141-1144, 1993

A/Title: The complete amino acid sequence of momordin-a, a ribosome-inactivating protein

A/Reference number: JN0628; MUID:93372485

A/Accession: JN0628

A/Molecule type: protein

A/Residues: 24-107, 'Q', '109-123, 125-147, 'U', '149-154, 'I', '156-205, 'I', '207-208, 'L', '210-21

A/Experimental source: seed

R/Reh, J.; Wang, Y.; Dong, Y.; Stuart, D.I.

submitted to the Brookhaven Protein Data Bank, January 1994

A/Reference number: A52272; PDB:1AHC

A/Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-269

R/Husain, J.; Tickle, I.J.; Wood, S.P.

submitted to the Brookhaven Protein Data Bank, March 1994

A/Reference number: A52385; PDB:1KOW

A/Contents: annotation; X-ray crystallography, 2.16 angstroms, residues 24-86, 'L', '88-

R/Huang, Q.; Liu, S.; Tang, Y.; Jin, S.; Wang, Y.

submitted to the Brookhaven Protein Data Bank, July 1994

A/Reference number: A67089; PDB:1MRH

A/Contents: annotation; X-ray crystallography, 2.0 angstroms, residues 24-77, 'R', '79-1

C/Function:

A/Description: hydrolyzes the N-glycosidic bond of a specific adenosine in 28S rRNA

C/Superfamily: RNA N-glycosidase; RNA N-glycosidase homology

C/Keywords: glycoprotein; glycosidase; hydrolase; lectin; seed; toxin

F/1-18/Domain: signal sequence #status predicted <SIG>

F/19-23/Domain: amino-terminal propeptide #status predicted <PRO>

F/24-269/Product: RNA N-glycosidase alpha-momorcharin #status experimental <MAT>

F/27-266/Domain: RNA N-glycosidase homology <RNG>

F/270-286/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F/93,183,186/Active site: Tyr, Glu, Arg #status predicted

F/250/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 1.9%; Score 9; DB 1; Length 286;

Best Local Similarity 100.0%; Pred. No. 0.92; Mismatches 0; Indels 0; Gaps 0;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 350 FLGSIYKG 358
|||||
Db 15 FLGSIYKG 23

RESULT 9

S56097 alpha-glucosidase - yeast (Saccharomyces fibuligera) (fragments)

C/Species: Saccharomyces fibuligera

C/Date: 10-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 07-May-1999

C/Accession: S56097

R/Reiser, V.; Gasperik, J.

Biochem. J. 308, 753-760, 1995

A/Title: Purification and characterization of the cell-wall-associated and extracellular

A/Reference number: S56097; MUID:97104271

A/Accession: S56097

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-85 <REL>

Query Match 1.7%; Score 8; DB 2; Length 85;

Best Local Similarity 100.0%; Pred. No. 3.3; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 429 VDTYKGL 436
|||||
Db 72 VDTYKGL 79

RESULT 10

H69813 hypothetical protein yfmQ - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C/Accession: H69813

R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertei, C.; Bron, S.; Brulliet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henuat, A.; Hilbert, H.; Holtsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinis, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel Y., M.; Ogawa, K.; Ogilawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetille Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadle, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Serot A.; Schuch, M.; Tamakoshi, A.; Tanaka, T.; Terpiltra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipal, A.; Yamamoto, K.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A:Reference number: A69580; MUID:98044033

A:Accession: H69813

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-148 <KUN>

A:Cross-references: GB:299107; GB:299108; GB:AL009126; NID:g2633055; PIDN:CAB12567.1; PI A:Experimental source: strain 168

C:Genetics:

A:Gene: yfmQ

C:Superfamily: *Bacillus subtilis* hypothetical protein yfmQ

Query Match 1.7%; Score 8; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 392 LSLDGRRL 399
|||||||
DB 46 LSLDGRRL 53

RESULT 11

H70248

hypothetical protein BBJ31 - Lyme disease spirochete plasmid J/lp38

C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: H70248

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White son, D.; Peterson, J.; Kervilave, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt, Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.

Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943

A:Accession: H70248

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-240 <KLE>

A:Cross-references: GB:AE000787; NID:g2690175; PIDN:AA66105.1; PID:g2690197; TIGR:BBJ31

C:Genetics:

A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

Query Match 1.7%; Score 8; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 152 LGDVKGNG 159
|||||||
DB 189 LGDVKGNG 196

RESULT 12

F70250

hypothetical protein BBJ45 - Lyme disease spirochete plasmid J/lp38

C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C:Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 08-Oct-1999

C:Accession: F70250

R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; Wh son, D.; Peterson, J.; Kervilave, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vu Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B. Nature 390, 580-586, 1997

A:Authors: Smith, H.O.; Venter, J.C.

A:Title: Genomic sequence of a Lyme disease spirochete, *Borrelia burgdorferi*.

A:Reference number: A70100; MUID:98065943

A:Accession: F70250

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-240 <KLE>

A:Cross-references: GB:AE000787; NID:g2690175; PIDN:AA66127.1; PID:g2690219; TIGR:BB A:Experimental source: strain B31

C:Genetics:

A:Genome: plasmid

Query Match 1.7%; Score 8; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 152 LGDVKGNG 159
|||||||
DB 189 LGDVKGNG 196

RESULT 13

T27695

hypothetical protein ZK1128.4 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27695

R:Berk, M.

submitted to the EMBL Data Library, January 1995

A:Reference number: Z20407

A:Accession: T27695

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-345 <WIL>

A:Cross-references: EMBL:Z47357; PIDN:CAA87423.1; GSPDB:GN00021; CESP:ZK1128.4

A:Experimental source: clone ZK1128

C:Genetics:

A:Gene: CESP:ZK1128.4

A:Map position: 3

A:Introns: 49/3; 83/3; 161/1; 267/3; 286/3

Query Match 1.7%; Score 8; DB 2; Length 345;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 450 PLGPALAH 457
|||||||
DB 150 PLGPALAH 157

RESULT 14

D95911

probable exported protein, slightly similar to protein involved in assembly of outer

C:Species: *Shimorhizobium meliloti*

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: D95911

R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb PSymb megaplasmid from the N2-fixing e

A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: D95911

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-429 <KUN>

A:Cross-references: GB:AL591985; PIDN:CAC48956.1; PID:g15140441; GSPDB:GN00167

A:Experimental source: strain 1021, megaplasmid PSymb

R;Gallier, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMB20816
 A:Genome: plasmid

Query Match 1.7%; Score 8; DB 2; Length 429;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 95 VLPSLISS 102
 |||||
 Db 21 VLPSLISS 28

RESULT 15
 T50687
 proline transport protein 3 [imported] - tomato
 C:Species: Lycopersicon esculentum (tomato)
 C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T50687
 R;Schwacke, R.; Grallath, S.; Breikreuz, K.E.; Stransky, E.; Stransky, H.; Frommer, W.B.
 Plant Cell 11, 377-392, 1999
 A:Title: LepTol1, a transporter for proline, glycine betaine, and gamma-amino butyric acid
 A:Reference number: 225181; MUID:99172053
 A:Accession: T50687
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-442 <SCH>
 A:Cross-references: EMBL:AF014810; PIDN:AAD25162.1
 C:Genetics:
 A:Gene: PROT3

Query Match 1.7%; Score 8; DB 2; Length 442;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 241 TLSIKDGL 248
 |||||
 Db 198 TLSIKDGL 205

Search completed: August 29, 2002, 16:52:53
 Job time: 147 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:52:26 ; Search time 13.42 Seconds
(Without alignments)
1361.820 Million cell updates/sec

Title: US-09-841-758-1

Perfect score: 472
Sequence: 1 MATKCGNGPGYSTPLFAMK.....PALAHLERLYPGGDSSDIWI 472

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	124	26.3	472	1 SBP1_HUMAN	Q13228 homo sapien
2	49	10.4	472	1 SBP1_MOUSE	P17563 mus musculu
3	43	9.1	472	1 SBP2_MOUSE	O63836 mus musculu
4	16	3.4	490	1 SBP_ARATH	O23264 arabidopsis
5	9	1.9	286	1 RIPL_MOMCH	P16094 momordica c
6	8	1.7	504	1 AMID_SUITO	P95896 sulfolobus
7	7	1.5	163	1 TCD2_ECOLI	P75916 escherichia
8	7	1.5	168	1 MENG_PASMU	O04907 caenorhabdi
9	7	1.5	212	1 EGGL_SCHJA	P19470 schistosoma
10	7	1.5	221	1 FEZ1_CORGL	O24748 corynebacte
11	7	1.5	221	1 MOVF_TOMIA	P29799 tomato mosa
12	7	1.5	264	1 ERMF_TOMIA	P10337 bacteroides
13	7	1.5	266	1 ERMF_BACFR	O02607 bacteroides
14	7	1.5	266	1 ERMF_BACFR	P18561 rhizobium m
15	7	1.5	326	1 SYRM_RHIME	P77269 escherichia
16	7	1.5	327	1 YPHF_ECOLI	O52812 rhizobium l
17	7	1.5	341	1 AAPJ_RHILV	P12719 cyanophora
18	7	1.5	360	1 AAOB_THRAC	P12719 cyanophora
19	7	1.5	360	1 PSBA_CYAPA	P22929 saccharomyc
20	7	1.5	380	1 CARP_SAFET	P40181 streptomyce
21	7	1.5	390	1 YGLI_STRCO	O98937 gallus gall
22	7	1.5	440	1 FXGA_CHICK	P34904 gallus gall
23	7	1.5	440	1 GACA_CHICK	P21265 gallus gall
24	7	1.5	459	1 PUR8_CHICK	P18508 rattus norv
25	7	1.5	466	1 GAC2_RAT	P18507 homo sapien
26	7	1.5	467	1 GAC2_HUMAN	P22733 mus musculu
27	7	1.5	474	1 GAC2_MOUSE	P22300 bos taurus
28	7	1.5	475	1 GAC2_BOVIN	P76658 escherichia
29	7	1.5	477	1 RFAE_ECOLI	P25313 azospirillum
30	7	1.5	482	1 NIFD_AZOB	P93338 n nado-depe
31	7	1.5	486	1 GAPN_NICPL	O92360 schizosacch
32	7	1.5	498	1 YDHF_SCHPO	P32206 swinepox vi
33	7	1.5	500	1 VC13_SPVKA	

34	7	1.5	511	1 MVP1_YEAST	P40959 saccharomyc
35	7	1.5	520	1 YEAE_SCHPO	O14079 schizosacch
36	7	1.5	537	1 SYE_TREPA	O83679 treponema p
37	7	1.5	560	1 YDZH_SCHPO	O10264 schizosacch
38	7	1.5	722	1 Z219_HUMAN	O9P2Y4 homo sapien
39	7	1.5	741	1 FIBA_CHICK	P14448 gallus gall
40	7	1.5	750	1 LEU2_RHINI	P55811 rhizopus m
41	7	1.5	857	1 PIFI_YEAST	P07271 saccharomyc
42	7	1.5	873	1 SYL_PSEAE	O9HX33 pseudomonas
43	7	1.5	1222	1 PM5P_HUMAN	O15155 homo sapien
44	7	1.5	1628	1 MAGR_CLOPE	P26831 clostridium
45	7	1.5	4486	1 DYH9_HUMAN	O9NYC9 homo sapien

ALIGNMENTS

RESULT 1	SBP1_HUMAN	STANDARD:	PRT:	472 AA.
AC	Q13228:			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Selenium-binding protein 1.			
GN	SELENBP1 OR SBP.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TRISUE-Heart;			
RA	MEDLINE-97179296; PubMed-9027582;			
RT	Chang P.W.G., Tsui S.K.W., Liew C., Lee C., Wayne M.M.Y., Fung K.;			
RT	"Isolation, characterization, and chromosomal mapping of a novel cDNA			
RT	clone encoding human selenium binding protein.";			
RL	J. Cell. Biochem. 64:217-224(1997).			
CC	- FUNCTION: NOT KNOWN; BIND SELENIUM.			
CC	- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.			
CC	- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL: U29091; AAB02395.1; -			
DR	MTM; 604188; -			
KW	Selenium.			
SQ	SEQUENCE 472 AA; 52313 MW; F484CF1CD68FC3B CRC64;			

Query Match 26.3%; Score 124; DB 1; Length 472;

Best Local Similarity 100.0%; Pred. No. 5.3e-119;

Matches 124; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	136	LHRSCHLASEVMISLGDYKGNKGKGFVLLDGETFEVKGTEWRPGGAAPLGDYFWYQPR	195
DB	136	LHRSCHLASEVMISLGDYKGNKGKGFVLLDGETFEVKGTEWRPGGAAPLGDYFWYQPR	195
QY	196	HNWISTEMAAPNVLRRGFNPADYEAGLYSHLYVMQREHYQYOTSLKDGILPLEIRF	255
DB	196	HNWISTEMAAPNVLRRGFNPADYEAGLYSHLYVMQREHYQYOTSLKDGILPLEIRF	255
QY	256	LHNP 259	
DB	256	LHNP 259	

```

RESULT 2
SBP1_MOUSE STANDARD: PRT: 472 AA.
AC P17563:
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Selenium-binding protein 1 (56 kDa selenium-binding protein) (SP56).
GN SELENBP1 OR LPSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91029855; PubMed=2225343;
RA Bansel M.P., Mukhopadhyay T., Scott J., Cook R.G., Mukhopadhyay R.,
RA Medina D.;
RT "DNA sequencing of a mouse liver protein that binds selenium:
RT implications for selenium's mechanism of action in cancer
RT prevention.";
RL Carcinogenesis 11:2071-2073(1990).
CC -I- FUNCTION: NOT KNOWN. BIND SELENIUM.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER, KIDNEY AND, TO A
CC LESSER EXTENT, LUNG.
CC -I- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: M32032; AAA40104.1; -.
DR PIR: S27878; S27878.
DR SWISS-2DPAGE: P17563; MOUSE.
DR MGD: MGI:96825; Selenbp1.
KW Selenium.
SQ SEQUENCE 472 AA; 52352 MW; D501292C4876033D CRC64;

Query Match
Best Local Similarity 10.4%; Score 49; DB 1; Length 472;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 CGPGYSTPLEAMKGPREEIVYLPCTYRNTGTAPDYLATVDVDPKSPQY 56
DB 8 CGPGYSTPLEAMKGPREEIVYLPCTYRNTGTAPDYLATVDVDPKSPQY 56

RESULT 3
SBP2_MOUSE STANDARD: PRT: 472 AA.
AC Q63836;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Selenium-binding protein 2 (56 kDa acetaminophen-binding protein)
DE (AP56).
GN SELENBP2 OR LPSB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=93201669; PubMed=8453708;
RA Lanfear J., Fleming J., Walker M., Harrison P.;

```

```

RT "Different patterns of regulation of the genes encoding the closely
RT related 56 kDa selenium- and acetaminophen-binding proteins in normal
RT tissues and during carcinogenesis.";
RL Carcinogenesis 14:335-340(1993).
-----
RP SEQUENCE OF 175-189; 196-220; 228-242; 290-296; 334-343 AND 399-408.
RX MEDLINE=92171951; PubMed=1540179;
RA Punford N.R., Martin B.M., Hinson J.A.;
RT "A metabolite of acetaminophen covalently binds to the 56 kDa
RT selenium binding protein.";
RL Biochem. Biophys. Res. Commun. 182:1348-1355(1992).
CC -I- FUNCTION: NOT KNOWN. BIND SELENIUM AND ACETAMINOPHEN.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic.
CC -I- TISSUE SPECIFICITY: MAINLY EXPRESSED IN LIVER.
CC -I- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: S56599; AAB25841.2; -.
DR MGD: MGI:104859; Selenbp2.
KW Selenium.
SQ SEQUENCE 472 AA; 52628 MW; C32FE819C4AD07CA CRC64;

Query Match
Best Local Similarity 9.1%; Score 43; DB 1; Length 472;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 TPLEAMKGPREEIVYLPCTYRNTGTAPDYLATVDVDPKSPQY 56
DB 14 TPLEAMKGPREEIVYLPCTYRNTGTAPDYLATVDVDPKSPQY 56

RESULT 4
SBP_ARATH STANDARD: PRT: 490 AA.
AC O23264;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative selenium-binding protein.
GN ATG414030 OR DL3055C.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=98121113; PubMed=9461215;
RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,
RA Bergkamp R., Dirkse W., van Staveren M., Stiekema W., Drost L.,
RA Ridley P., Hudson S.-A., Patel K., Murphy G., Piffanelli P.,
RA Wedler H., Wedler E., Wambutt R., Weitzenegger T., Pohl T.M.,
RA Terryn N., Gieles J., Villariol R., de Clerck R., van Montagu M.,
RA Lecharny A., Auborg S., Gy I., Kreis M., Lao N., Kavanagh T.,
RA Hempel S., Kotter P., Entian K.-D., Rieger M., Schaeffer M.,
RA Mueller-Auer S., Silvey M., James R., Montfort A., Pons A.,
RA Puigdomenech P., Douka A., Vouklatou E., Milioni D., Hatzopoulos P.,
RA Piravandi E., Obermaier B., Hilbert H., Duesterhoef A., Moore T.,
RA Jones J.D.G., Eneva T., Palme K., Benes V., Reichman S., Ansoore W.,
RA Cooke R., Berger C., Delsey M., Voelt M., Volckaert G., Mewes H.-W.,
RA Klosterman S., Schueller C., Chaiwatsis N.;
RT "Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of
RT Arabidopsis thaliana.";
RL Nature 391:485-488(1998).

```

[2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083486; PubMed=10617196;
 RA Meyer K.F.X., Schueller C., Wandut R., Murphy G., Volckaert G.,
 RA Pohl T., Duesterhoeft A., Stiekema W., Entlan K.-D., Terryn N.,
 RA Harris B., Ansoege W., Brandt P., Grivell L.A., Rieger M.,
 RA Weichselgartner M., de Simone V., Obermaier B., Mache R., Mueller M.,
 RA Kreis M., Delseny M., Puigdomenech P., Watson M., Schmitthenel T.,
 RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 RA Vos P., Hohnel J., Zimmermann W., Wedler J., Robben J.,
 RA Langham S.-A., McCullagh B., Bilham L., Robben J.,
 RA van der Schueren J., Gymonprez B., Chuang Y.-J., Vandebussche F.,
 RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Melzenger T., Bothe G., Rampsperger U., Hilbert H., Braum M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dikse W.,
 RA Woolman P., Klein Lankhorst R., Rose M., Haut J., Koelter P.,
 RA Bernsner S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,
 RA De Keyser A., Buyshaert C., Gielen J., Villarroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Leonard N., Mclay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Blocker H., Scharte M., Grimm M., Loehner T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Neumann S., Argiridou A., Vitale D., Liguori R., Piravandi E.,
 RA Massenet O., Quigley F., Clapaud G., Muendlein A., Felber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Cheford F., Cooke R., Berger C., Monfort A., Cascuberta E.,
 RA Glibons T., Weber N., Vandenbol M., Barges M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 RA Fishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedina N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abul-Friedeh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Lattelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Mux P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pegin K., Hillier L.,
 RA Nelson J., Spieh J., Ryan E., Andrews S., Giesel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Saeby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hamed A., Lohi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.;
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 thaliana.";
 RL Nature 402:769-777(1999).
 CC -1- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; Z97335; CAB10182.1; -;
 DR EMBL; AL161537; CAB78445.1; -;
 KW Selenium.
 SQ SEQUENCE 490 AA; 54057 MW; 10EE7B9BCF2P0390 CRC64;

RESULT 5
 RIPI_MOMCH STANDARD; PRT; 286 AA.
 ID RIPI_MOMCH
 AC P16094; P24697;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Ribosome-inactivating protein momordin I precursor (rRNA
 DE N-glycosidase) (EC 3.2.2.22) (Alpha-momorcharin) (Alpha-MMC).
 OS Momordica charantia (Bitter melon) (Balsam pear).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucotsids I; Cucurbitales; Cucurbitaceae; Momordica.
 OC NCBI_TaxID=3673;
 OX
 RN
 RN SEQUENCE FROM N.A.
 RP TISSUE=Seed;
 RC MEDLINE=91159486; PubMed=2001404;
 RX Ho W.K.K., Liu S.C., Shaw P.C., Yeung H.W., Ng T.B., Chan W.Y.;
 RA "Cloning of the cDNA of alpha-momorcharin: a ribosome inactivating
 RT protein.";
 RL Biochim. Biophys. Acta 1088:311-314(1991).
 RN [2]
 RP SEQUENCE OF 24-38.
 RC TISSUE=Seed;
 RX MEDLINE=89326691; PubMed=2753596;
 RA Montecucchi P.C., Lazzarini A.M., Barbieri L., Stirpe F., Soria M.,
 RA Leppl D.;
 RT "N-terminal sequence of some ribosome-inactivating proteins.";
 RL Int. J. Pept. Protein Res. 33:263-267(1989).
 RN [3]
 RP SEQUENCE OF 24-70.
 RC TISSUE=Seed;
 RX MEDLINE=89005108; PubMed=3262509;
 RA Casellas P., Dussossoy D., Falasca A.I., Barbieri L.,
 RA Guillemot J.C., Ferrara P., Bolognesi A., Genini P., Stirpe F.;
 RT "Trichostatin, a ribosome-inactivating protein from the seeds of
 RT Trichostema kirilowii Maximowicz. Purification, partial
 RT characterization and use for preparation of immunotoxins.";
 RL Eur. J. Biochem. 176:581-588(1988).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
 RX MEDLINE=94356447; PubMed=8075985;
 RA Ren J., Wang Y., Dong Y., Stuart D.I.;
 RT "The N-glycosidase mechanism of ribosome-inactivating proteins
 RT implied by crystal structures of alpha-momorcharin.";
 RL Structure 2:7-16(1994).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (2.16 ANGSTROMS).
 RX MEDLINE=94192822; PubMed=8143869;
 RA Husain J., Tickle I.J., Wood S.P.;
 RT "Crystal structure of momordin, a type I ribosome inactivating
 RT protein from the seeds of Momordica charantia.";
 RL PNAS Lett. 342:154-158(1994).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=95344383; PubMed=7619070;
 RA Huang Q., Liu S., Tang Y., Jin S., Wang Y.;
 RT "Studies on crystal structures, active-centre geometry and
 RT depurinating mechanism of two ribosome-inactivating proteins.";
 RL Biochem. J. 309:285-298(1995).
 CC -1- CATALYTIC ACTIVITY: Endonucleolysis of the N-glycosidic bond at one
 CC specific adenosine on the 28S rRNA.
 CC
 CC -1- SIMILARITY: BELONGS TO THE RIBOSOME-INACTIVATING FAMILY. TYPE 1
 CC RIP SUBFAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL; X57682; CAA40869.1; -

DR PIR; S14273; RLPUG.

DR PIR; S16490; S16490.

DR PDB; 1AHB; 22-JUN-94.

DR PDB; 1AHB; 22-JUN-94.

DR PDB; 1AHC; 22-JUN-94.

DR PDB; 1MOM; 31-MAY-94.

DR PDB; 1MRG; 07-FEB-95.

DR PDB; 1MRH; 07-FEB-95.

DR PDB; 1MRI; 07-FEB-95.

DR GlycosylatedB; P16094; -

DR InterPro; IPR001574; RIP.

DR Pfam; PF00161; RIP; 1.

DR PRINTS; PR00396; SHIGARICIN.

DR PROSITE; PS00275; SHIGA.RICIN; 1.

KM Protein synthesis inhibitor; Hydrolase; Toxin; Signal; Glycoprotein;

KW 3D-structure.

FT SIGNAL 1 23

FT CHAIN 24 269 RIBOSOME-INACTIVATING PROTEIN MOMORDIN I.

FT PROPEP 270 286 MISSING IN MATURE PROTEIN.

FT ACT_SITE 183 183

FT CAROHD 250 250 N-LINKED (GLCNAc. . .).

FT SEQUENCE 286 AA; 31532 MW; E1B013ABEBC216CF CRC64;

Query Match 1.9%; Score 9; DB 1; Length 286;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 350 FLGGSIVKG 358
| | | | | | | | | |
DB 15 FLGGSIVKG 23

RESULT 6
ID AMID_SULISO STANDARD; PRT; 504 AA.
AC P95896;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Amidase (EC 3.5.1.4).
GN SS02122 OR C02016 OR C02.017.
OS Sulfolobus solfataricus.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
ON NCBI_TaxID=2287;
RX [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC STRAIN-DSM 5833 / MT-4;
RX MEDLINE=21346015; PubMed=11453462;
RA d'Abusco A.S., Amendola S., Scandurra R., Politi L.;
RT "Molecular and biochemical characterization of the recombinant amidase from hyperthermophilic archaeon Sulfolobus solfataricus";
RL Extremophiles 5:183-192(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=97053432; PubMed=8899719;
RA Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,
RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,
RA Doolittle W.F., Ragan M.A., Charlebois R.L.;
RT "Organizational characteristics and information content of an archaeal genome: 156 kb of sequence from Sulfolobus solfataricus
RT P2.";
RL Mol. Microbiol. 22:175-191(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35092 / DSM 1617 / P2;
RX MEDLINE=21332296; PubMed=11427726;
RA

RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Awez M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Ehras G., Fletcher C., Gordon P.M.K.,
RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -!- FUNCTION: Enantioselective, active on 2-to 6-carbon aliphatic
CC amides and on many aromatic amides.
CC -!- CATALYTIC ACTIVITY: A monocarboxylic acid amide + H(2)O -> a
CC monocarboxylate + NH(3).
CC -!- SUBUNIT: Monomer.
CC -!- MISCELLANEOUS: Active over the pH range 4-9 and at temperatures
CC from 60 degrees to 95 degrees Celsius.
CC -!- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; AF290611; AAK83092.1; -

DR EMBL; Y08256; CAA69448.1; -

DR EMBL; AE006819; AAK42301.1; -

DR InterPro; IPR000120; Amidase.

DR Pfam; PF01425; Amidase; 1.

DR PROSITE; PS00571; AMIDASES; 1.

KW Hydrolase; Complete proteome.

SEQUENCE 504 AA; 55655 MW; A9103AB7D09A88EE CRC64;

Query Match 1.7%; Score 8; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 364 LDEDELKS 371
| | | | | | | | | |
DB 22 LDEDELKS 29

RESULT 7
ID YCDZ_ECOLI STANDARD; PRT; 163 AA.
AC P75916;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ycdz.
GN YCDZ OR B1036.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
ON NCBI_TaxID=562;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aliba H., Baba T., Fujita K., Hayashi K., Honjo A.,

RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.,
RT A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
corresponding to the 12,7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO E.COLI YAHC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000205; AAC74120.1; ALT_INIT.
DR EMBL: D90740; BAA35817.1; ALT_INIT.
DR EMBL: D90741; BAA35826.1; ALT_INIT.
DR Ecogene; EG3872; ycdZ.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
SQ SEQUENCE 163 AA; 17074 MW; 999B93A1FE35B68 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 163;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 93 KLVLPSTL 99
Db 122 KLVLPSTL 128

RESULT 8
MENG_PASMU
ID MENG_PASMU STANDARD; PRT; 166 AA.
AC O9CLP9;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE S-adenosylmethionine:2-demethylmenaquinone methyltransferase
(EC 2.1.1.1).
GN MENG OR PM1168
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OX Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=2115866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genome sequence of *Pasteurella multocida* PM70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- FUNCTION: CONVERTS DIMETHYLMENAQUINONE (DMK) TO MENAQUINONE (MK)
(By similarity).
CC -1- PATHWAY: LAST STEP IN MENAQUINONE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE MENG FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006157; AK03252.1; -
KW Menaquinone biosynthesis; Transferase; Methyltransferase;
KM Complete proteome.
SQ SEQUENCE 166 AA; 18144 MW; 63A19ED83429CB1 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 166;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 308 GLTRDIL 314
Db 46 GLTRDIL 52

RESULT 9
YIAL_CAEEL
ID YIAL_CAEEL STANDARD; PRT; 168 AA.
AC Q04907;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Hypothetical 19.1 kDa protein C30C11.1 in chromosome III.
GN C30C11.1
OS *Caenorhabditis elegans*.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; *Caenorhabditis*.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Wohlschlag P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of *C.*
RT *elegans*.";
RL Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L09634; AAA27965.1; -
DR PIR: S44782; S44782.
DR WormPep: C30C11.1; CE00100.
KW Hypothetical protein
SQ SEQUENCE 168 AA; 19107 MW; 9B01DBAD6979C4E5 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 168;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 377 VVGKRV 383
Db 144 VVGKRV 150

RESULT 10
 EGG1_SCHJA STANDARD: PRT: 212 AA.
 ID EGG1_SCHJA
 AC P19470:
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DE Eggshell protein 1 precursor.
 GN EGG-1.
 OS Schistosoma japonicum (Blood fluke).
 CC Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminths;
 CC Rhabdiorhiza; Eulicthophora; Reversopectata; Mediofusata;
 CC Neodermata; Trematoda; Digenea; Strigoida; Schistosomatidae;
 CC Schistosomatidae; Schistosoma.
 RX NCBI_TaxID=6182;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=91042837; PubMed=2172818;
 RA Henkle K.J., Cook G.A., Foster L.A., Engman D.M., Bobek L.A.,
 RA Cain G.D., Donelson J.E.;
 RT "The gene family encoding eggshell proteins of Schistosoma
 japonicum";
 RT Mol. Biochem. Parasitol. 42:69-82(1990).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: M32281; AAA29875.1; -
 DR PIR: A44994; A44994.
 DR HSSP: P30129; ADPV.
 DR InterPro: IPR002952; Eggshell.
 DR PRINTS: PR01228; EGGSHELL.
 DR Multigene family; Eggshell; Repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 212 EGGSHELL PROTEIN 1.
 SQ SEQUENCE 212 AA; 18937 MW; 32059D27FDA21FCB CRC64;

Query Match 1.5%; Score 7; DB 1; Length 212;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 156 KNGKGG 162
 DB 183 KNGKGG 189

RESULT 11
 YFZ1_CORGL STANDARD: PRT: 221 AA.
 ID YFZ1_CORGL
 AC O24748:
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical 24.0 kDa protein in FTSZ 3 region.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
 CC Corynebacterium.
 RX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97382442; PubMed=9240446;
 RA KODAYASHI M., ASAI Y., HATAKEYAMA K., KIJIMA N., WACHI M., NAGAI K.,
 RA YUKAWA H.;
 RT "Cloning, sequencing, and characterization of the ftsz gene from
 RT corynebacterium bacteria";

RL Biochem. Biophys. Res. Commun. 236:383-388(1997).
 CC -1- SIMILARITY: BELONGS TO THE UPF0001 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AB003132; BAA21689.1; -
 DR InterPro: IPR001608; UPF0001.
 DR Pfam: PF01168; UPF0001; 1.
 DR PROSITE: PS01211; UPF0001; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 221 AA; 24029 MW; DD8DD041632B2C9D CRC64;

Query Match 1.5%; Score 7; DB 1; Length 221;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 390 IQLSLDG 396
 DB 119 IQLSLDG 125

RESULT 12
 MOV_P_TOMLA STANDARD: PRT: 264 AA.
 ID MOV_P_TOMLA
 AC P29799:
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Movement protein (cell-to-cell transport protein) (30 kDa protein).
 GN MP.
 OS Tomato mosaic virus (strain Lila) (TOMV).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
 RX NCBI_TaxID=31748;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92113565; PubMed=1730937;
 RA CALDER V.L., PALUKAITIS P.;
 RT "Nucleotide sequence analysis of the movement genes of resistance
 RT breaking strains of tomato mosaic virus.";
 RL J. Gen. Virol. 73:165-168(1992).
 CC -1- FUNCTION: INVOLVED IN TRANSPORT OF THE VIRUS FROM THE INITIALLY
 CC INFECTED CELLS TO ADJACENT CELLS. POSSIBLY BY MODIFYING THE
 CC FUNCTION OF THE PLASMODESMATA. ALSO INFLUENCES LOCAL LESION
 CC DEVELOPMENT. BINDS TO SINGLE-STRANDED NUCLEIC ACID.
 CC -1- SIMILARITY: BELONGS TO THE TOBAMOVIRUSES MOVEMENT PROTEIN FAMILY.
 DR PIR: J01456; MMBVL1.
 DR InterPro: IPR001022; Tobamo_MP.
 DR Pfam: PF01107; Tobamo_MP; 1.
 DR PRINTS: PR00964; MOVEMENT.
 KW DNA-binding; Transport.
 SQ SEQUENCE 264 AA; 29354 MW; 16E9675B6680F391 CRC64;

Query Match 1.5%; Score 7; DB 1; Length 264;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 375 PLVYKGG 381
 DB 2 PLVYKGG 8

RESULT 13
 ERMF_BACFR STANDARD: PRT: 266 AA.
 ID ERMF_BACFR
 AC P10337;


```

DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE rRNA adenine N-6-methyltransferase (EC 2.1.1.48) (Macrolide-
DE lincosamide-streptogramin B resistance protein) (Erythromycin
DE resistance protein).
GN ermF.
OS Bacteroides fragilis.
OG Plasmid pBFA.
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
OC Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-TnA351;
RX MEDLINE=87056929; PubMed=3023281;
RA Rasmussen J.L., Odelson D.A., Macrina F.L.;
RT "Complete nucleotide sequence and transcription of ermF, a macrolide-
RT lincosamide-streptogramin B resistance determinant from Bacteroides
RT fragilis."
RL J. Bacteriol. 168:523-533(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-TnA351;
RX MEDLINE=87279905; PubMed=3038844;
RA Rasmussen J.L., Odelson D.A., Macrina F.L.;
RT "Complete nucleotide sequence of insertion element IS4351 from
RT Bacteroides fragilis."
RL J. Bacteriol. 169:3573-3580(1987).
CC -1- FUNCTION: THIS PROTEIN PRODUCES A DIMETHYLATION OF THE ADENINE
CC RESIDUE AT POSITION 2058 IN 23S RNA, RESULTING IN REDUCED
CC AFFINITY BETWEEN RIBOSOMES & MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B
CC ANTIBIOTICS.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing N6-methyladenine.
CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14730; AAA96217.1; -
DR EMBL; M17124; AAA88675.1; -
DR PIR; A25157; A25157.
DR HSSP; P13956; IQAM.
DR InterPro: IPR001737; RNA_A_dimeth.
DR InterPro: IPR000051; SAM_bind.
DR Pfam; PF00398; RnaAD; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Transferase; Methyltransferase; Plasmid;
KW Transposable element.
SQ SEQUENCE 266 AA; 30356 MW; C7822A5D978C6274 CRC64;

```

```

Query Match 1.5%; Score 7; DB 1; Length 266;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 350 FLAGSTIV 356
DB 124 FLAGSTIV 130

```

```

RESULT 14
ID ERMU_BACFR STANDARD; PRT; 266 AA.
AC Q02607;
DT 01-JUL-1993 (Rel. 26, Created)

```

```

DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE rRNA adenine N-6-methyltransferase (EC 2.1.1.48) (Macrolide-
DE lincosamide-streptogramin B resistance protein) (Erythromycin
DE resistance protein).
GN ermF.
OS Bacteroides fragilis.
OG Plasmid V503.
OC Bacteria; CFB group; Bacteroidetes; Bacteroidales; Bacteroidaceae;
OC Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91288232; PubMed=1905805;
RA Halula M., Manning S., Macrina F.L.;
RT "Nucleotide sequence of ermF, a macrolide-lincosamide-streptogramin
RT (MLS) resistance gene encoding an RNA methylase from the conjugal
RT element of Bacteroides fragilis V503."
RL Nucleic Acids Res. 19:3453-3453(1991).
CC -1- FUNCTION: THIS PROTEIN PRODUCES A DIMETHYLATION OF THE ADENINE
CC RESIDUE AT POSITION 2058 IN 23S RNA, RESULTING IN REDUCED
CC AFFINITY BETWEEN RIBOSOMES & MACROLIDE-LINCOSAMIDE-STREPTOGRAMIN B
CC ANTIBIOTICS.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + rRNA = S-adenosyl-L-
CC homocysteine + rRNA containing N6-methyladenine.
CC -1- SIMILARITY: BELONGS TO THE RNA ADENINE N-6-METHYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62487; AAA63165.1; -
DR PIR; S34413; S34413.
DR InterPro: IPR001737; RNA_A_dimeth.
DR InterPro: IPR000051; SAM_bind.
DR Pfam; PF00398; RnaAD; 1.
DR PROSITE; PS01131; RNA_A_DIMETH; 1.
KW Antibiotic resistance; Transferase; Methyltransferase; Plasmid;
KW SEQUENCE 266 AA; 30424 MW; B9F2ABDAB1AF9E00 CRC64;

```

```

Query Match 1.5%; Score 7; DB 1; Length 266;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 350 FLAGSTIV 356
DB 124 FLAGSTIV 130

```

```

RESULT 15
ID SYRM_RHIME STANDARD; PRT; 326 AA.
AC P18561; Q53001;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE SYRM protein (Symbiotic regulator).
GN SYRM OR RA0462 OR SMA0849.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA (megaplasmid 1).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / S047;
RX MEDLINE=90299788; PubMed=2361944;

```

```

RA Barnett M.J., Long S.R.;
RT "DNA sequence and translational product of a new nodulation-regulatory
RT locus: sym has sequence similarity to NodD proteins.";
RL J. Bacteriol. 172:3695-3700(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-AK631;
RX MEDLINE-92236421; PubMed-1809842;
RA Kondorosi E., Buitre M., Cren M., Iyer N., Hoffmann B., Kondorosi A.;
RT "Involvement of the sym and nodD genes of Rhizobium meliloti in nod
RT gene activation and in optimal nodulation of the plant host.";
RL Mol. Microbiol. 5:3035-3048(1991).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-1021;
RX MEDLINE-21396509; PubMed-11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubier F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kaiman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSyma megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -!- FUNCTION: SYRM ACTS IN TRANS TO STIMULATE NOD GENE EXPRESSION VIA
CC NODD3 AND EXO GENE EXPRESSION VIA SYRA.
CC -!- MISCELLANEOUS: THE SYRM ORF POSSESSED TWO TRANSLATION START CODONS
CC IN PROXIMITY TO EACH OTHER LEADING TO A PROTEIN OF 326 OR 301 AA.
CC -!- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33495; AAB38371.1; -
DR EMBL; X61396; CAA43667.1; ALT_INIT.
DR EMBL; AE007236; AAK65120.1; -
DR PIR; A44505; A44505.
DR InterPro; IPR000847; HTH_LYSR.
DR InterPro; IPR001583; NodD_C.
DR Pfam; PF00126; HTH_1; 1.
DR Pfam; PF01046; NodD_C_term; 1.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW Nodulation; Transcription regulation; DNA-binding; Activator; Plasmid;
KW Complete proteome.
FT DNA_BIND 49 68 H-T-H MOTIF (BY SIMILARITY).
FT CONFLICT 9 24 PHRAKAGVSDAAQOR -> RIGPNLPSVTPHNK (IN
FT REF. 2).
FT CONFLICT 227 227 D -> G (IN REF. 2).
SQ SEQUENCE 326 AA: 35897 MW: 412F888B1FF6C1 CRC64;

```

```

Query Match 1.5%; Score 7; DB 1; Length 326;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 182 GAAPLGY 188
DB 178 GAAPLGY 184

```

Search completed: August 29, 2002, 16:56:30
Job time: 244 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:52:06 ; Search time 30.95 Seconds
(without alignments)
2638.243 Million cell updates/sec

Title: US-09-841-758-1

Sequence: 1 MATKCGNCGPGYSTPLEAMK.....PALAHLATVPDGDSSDIWI 472

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp._archaea:*
2: sp._bacteria:*
3: sp._fungi:*
4: sp._human:*
5: sp._invertebrate:*
6: sp._mammal:*
7: sp._mhc:*
8: sp._organelle:*
9: sp._phage:*
10: sp._plant:*
11: sp._rodent:*
12: sp._virus:*
13: sp._vertebrate:*
14: sp._unclassified:*
15: sp._virus:*
16: sp._bacteriophage:*
17: sp._archaeoprotein:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	472	100.0	472	4	096GX7
2	217	46.0	472	4	09H8A8
3	49	10.4	472	11	091X87
4	39	8.3	84	11	09R0W8
5	39	8.3	84	11	09R0W7
6	16	3.4	470	10	093WU0
7	16	3.4	488	10	093WS1
8	15	3.2	488	13	09DPE9
9	15	3.2	478	10	023265
10	15	3.2	487	10	09A5S5
11	15	3.2	487	10	09A5S5
12	13	3.0	457	10	09AVW6
13	13	2.8	478	10	093YH7
14	11	2.3	146	10	042178
15	11	2.3	487	10	093VA1
16	10	2.1	468	5	09VFE24

17	10	2.1	471	5	09XXF9	09xxf9 caenorhabdi
18	10	2.1	480	10	09LK38	09lk38 arabidopsis
19	10	2.1	487	5	044507	044507 caenorhabdi
20	8	1.7	77	12	037004	037004 newcastle d
21	8	1.7	77	12	09DK19	09dk19 newcastle d
22	8	1.7	148	16	006475	006475 bacillus su
23	23	1.7	240	16	050786	050786 borrelia bu
24	24	1.7	240	16	050800	050800 borrelia bu
25	25	1.7	345	5	009645	009645 caenorhabdi
26	26	1.7	364	12	09WG26	09wg26 newcastle d
27	27	1.7	364	12	09WG28	09wg28 newcastle d
28	8	1.7	364	12	0914X4	0914x4 newcastle d
29	8	1.7	429	16	092VZ1	092vz1 rhizobium m
30	30	1.7	442	10	09XE50	09xe50 lycopersico
31	8	1.7	702	10	09FTD8	09ftd8 oryza sativ
32	32	1.7	3419	11	055147	055147 ratius norv
33	33	1.5	55	6	09TPE3	09tpe3 pithacia pi
34	34	1.5	64	10	042227	042227 arabidopsis
35	35	1.5	72	5	023632	023632 caenorhabdi
36	36	1.5	72	5	09XV78	09xv78 caenorhabdi
37	37	1.5	77	17	074107	074107 pyrococcus
38	38	1.5	88	2	09ZET6	09zet6 xanthobacte
39	39	1.5	126	8	09T228	09t228 phytophlor
40	40	1.5	132	16	09RW04	09rw04 delnoccocus
41	133	1.5	133	3	008931	008931 saccharomyc
42	41	1.5	145	16	093213	093213 staphylococ
43	43	1.5	153	10	082463	082463 nicotiana c
44	7	1.5	153	10	09XFE9	09xfe9 medicago sa
45	7	1.5	166	10	09C5T5	09c5t5 arabidopsis

ALIGNMENTS

RESULT 1
ID 096GX7 PRELIMINARY; PRT; 472 AA.
AC 096GX7;
DT 01-DEC-2001 (TREMUREL. 19, Created)
DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
DE SIMILAR TO SELENIUM BINDING PROTEIN 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON ADENOCARCINOMA;
RA Strausberg R.;
DR EMBL; BC009084; AA009084.1; -
SQ SEQUENCE 472 AA; 52391 MW; 6DC68F9B45FEC1BC CRC64;

Query Match 100.0%; Score 472; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKCGNCGPGYSTPLEAMKPREIYVLPICYNNTGEADYLAIVDVPKSPQYCOVI 60
DB 1 MATKCGNCGPGYSTPLEAMKPREIYVLPICYNNTGEADYLAIVDVPKSPQYCOVI 60
QY 61 HRLPMVNLKDELHSHGNNQSCSCGDSSTKSTKTVLPSLISRYVVDVSGEPAPLTK 120
DB 61 HRLPMVNLKDELHSHGNNQSCSCGDSSTKSTKTVLPSLISRYVVDVSGEPAPLTK 120
QY 121 VIEPKDIHAKCELAFITSHCLASGEVYMSISLGVKNGKGFVLDGETFEVVGTERP 180
DB 121 VIEPKDIHAKCELAFITSHCLASGEVYMSISLGVKNGKGFVLDGETFEVVGTERP 180
QY 181 GGAAPLDYDEWYQPRHNVMISTEWAAPNVLDGEPADVEAGLYGSHLYWDMORHEIVQ 240
DB 181 GGAAPLDYDEWYQPRHNVMISTEWAAPNVLDGEPADVEAGLYGSHLYWDMORHEIVQ 240

DB 181 GGAAPLGYDFWYQPRHNWISTEWAAPNVLRDGFNPADVEAGLYGSHLYVMDQRHEIVQ 240
 QY 241 TLTKDGLIPIELIRFLHNPDAAGFGVCAASSTIQRFYKNEGTSVEKVIQVPPKKVG 300
 DB 241 TLTKDGLIPIELIRFLHNPDAAGFGVCAASSTIQRFYKNEGTSVEKVIQVPPKKVG 300
 QY 301 WLLPEKGLITDILSLDRFLYFSNWLHGDLROYDISPPQPRRLTGQLFLGSIYKGP 360
 DB 301 WLLPEKGLITDILSLDRFLYFSNWLHGDLROYDISPPQPRRLTGQLFLGSIYKGP 360
 QY 361 VOVLDEDELKSOPEPLVYVKGKRVAGGPMIOLSLDGKRLYITTSIYSADKQFYPLIRE 420
 DB 361 VOVLDEDELKSOPEPLVYVKGKRVAGGPMIOLSLDGKRLYITTSIYSADKQFYPLIRE 420
 QY 421 GSVMLQVDVDTYKGGKLNPNFLVDFGKEPLPALAHELRYPGDCSSDIWI 472
 DB 421 GSVMLQVDVDTYKGGKLNPNFLVDFGKEPLPALAHELRYPGDCSSDIWI 472

RESULT 2

Q9H8A8 PRELIMINARY; PRT; 235 AA.
 AC Q9H8A8:
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE CDNA FLJ13813 FIS, CLONE THYROID00358, MODERATELY SIMILAR TO
 DE SELENIUM-BINDING LIVER PROTEIN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=THYROID;
 RA Itoigai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
 RA Arima M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.,
 RT "NEO human cDNA sequencing project";
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL, AK023875; BAB14709.1; -;
 DR InterPro: IPR000531; TonB_Box.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
 SQ SEQUENCE 235 AA; 25953 MW; D68FF73A616D986 CRC64;

Query Match 46.0%; Score 217; DB 4; Length 235;
 Best Local Similarity 100.0%; Pred. No. 2.4e-219;
 Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 MNLDELHSGMNTCCSFGDSTKSRTKLVPISLISSRIYVVDGSEPRAKLHVIEP 124
 DB 1 MNLDELHSGMNTCCSFGDSTKSRTKLVPISLISSRIYVVDGSEPRAKLHVIEP 60
 QY 125 KDIIHAKCELAFHTSHCLASGEVMISSLDGVKNGKGVLLDGETFEFKGTWERPGCA 184
 DB 61 KDIIHAKCELAFHTSHCLASGEVMISSLDGVKNGKGVLLDGETFEFKGTWERPGCA 120
 QY 185 PLGYDFWYQPRHNWISTEWAAPNVLRDGFNPADVEAGLYGSHLYVMDQRHEIVQ 244
 DB 121 PLGYDFWYQPRHNWISTEWAAPNVLRDGFNPADVEAGLYGSHLYVMDQRHEIVQ 180
 QY 245 KGLIPIELIRFLHNPDAAGFGVCAASSTIQRFYKNEGTSVEKVIQVPPKKVG 281
 DB 181 KGLIPIELIRFLHNPDAAGFGVCAASSTIQRFYKNEGTSVEKVIQVPPKKVG 217

RESULT 3
 Q91X87 PRELIMINARY; PRT; 472 AA.
 AC Q91X87;
 Q91X87;

DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE UNKNOWN (PROTEIN FOR MGC:18519).
 OS Mus musculus (mouse)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC011202; AAH11202.1; -;
 SQ SEQUENCE 472 AA; 52514 MW; 099BCE085668DE36 CRC64;

Query Match 10.4%; Score 49; DB 11; Length 472;
 Best Local Similarity 100.0%; Pred. No. 1.7e-42;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 CGPGSTPLEAKGPREETIVLPCTIRNTGTAPYLATVDVDPKSPQY 56
 DB 8 CGPGSTPLEAKGPREETIVLPCTIRNTGTAPYLATVDVDPKSPQY 56

RESULT 4

Q9ROW8 PRELIMINARY; PRT; 84 AA.
 AC Q9ROW8:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE SELENIUM LIVER BINDING PROTEIN SP56 (SLP-56).
 GN LPSB1.
 OS Spretus (Western wild mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10096;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPAIN (F26);
 RA Dregan T.A.;
 RL Submitted (Jan-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL, X95321; CAA64628.1; -;
 FT VARIANT 20 A -> D.
 FT VARIANT 35 M -> V.
 FT VARIANT 39 I -> V.
 SQ SEQUENCE 84 AA; 9408 MW; 77CCFCA4B535EB9 CRC64;

Query Match 8.3%; Score 39; DB 11; Length 84;
 Best Local Similarity 100.0%; Pred. No. 1.1e-32;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 434 GGLKLNPNFLVDFGKEPLPALAHELRYPGDCSSDIWI 472
 DB 46 GGLKLNPNFLVDFGKEPLPALAHELRYPGDCSSDIWI 84

RESULT 5
 Q9ROW7 PRELIMINARY; PRT; 84 AA.
 AC Q9ROW7:
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)
 DE SELENIUM LIVER BINDING PROTEIN (AP-56).
 GN LPSB2.
 OS Mus spretus (Western wild mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10096;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SPAIN (F26);
 RA Dragani T.A.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: X95322; CAA64629.1; -
 SQ SEQUENCE 84 AA; 9406 MW; B17B9CA3DC2A761B CRC64;

Query Match 8.3%; Score 39; DB 11; Length 84;
 Best Local Similarity 100.0%; Pred. No. 1.1e-32;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 434 GGLKLNPNFLVDEGKEPLGPAALHELRYPGDCSSDIWI 472
 |||||
 Db 46 GGLKLNPNFLVDEGKEPLGPAALHELRYPGDCSSDIWI 84

RESULT 6
 Q93WU0 ID Q93WU0 PRELIMINARY; PRT; 470 AA.
 AC Q93WU0;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE SLENIUM BINDING PROTEIN (FRAGMENT).
 GN SLP.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=FLOWER;
 RA Fietmekakis E.;
 RT "A novel Lotus japonicus nodulin gene is highly conserved among plants
 and animals, and encodes a homologue to the mammalian selenium binding
 protein.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ401229; CAC67446.1; -
 FT NON_TER 1
 SQ SEQUENCE 470 AA; 52109 MW; CB77FC3617F9634D CRC64;

Query Match 3.4%; Score 16; DB 10; Length 470;
 Best Local Similarity 100.0%; Pred. No. 7.9e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 385 GGPQMIOQLSDGKRLY 400
 |||||
 Db 383 GGPQMIOQLSDGKRLY 398

RESULT 7
 Q93WS1 ID Q93WS1 PRELIMINARY; PRT; 488 AA.
 AC Q93WS1;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE SLENIUM BINDING PROTEIN.
 GN SLP.
 OS Medicago sativa (Alfalfa).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 OX NCBI_TaxID=3879;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=SOMATIC EMBRYOS;
 RA Fietmekakis E.;
 RT "A novel Lotus japonicus nodulin gene is highly conserved among plants

RT and animals, and encodes a homologue to the mammalian selenium binding
 protein.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ401228; CAC67501.1; -
 SQ SEQUENCE 488 AA; 54109 MW; D0BD0E38D3F50B5D CRC64;

Query Match 3.4%; Score 16; DB 10; Length 488;
 Best Local Similarity 100.0%; Pred. No. 8.2e-08;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 385 GGPQMIOQLSDGKRLY 400
 |||||
 Db 401 GGPQMIOQLSDGKRLY 416

RESULT 8
 Q9DFK9 ID Q9DFK9 PRELIMINARY; PRT; 68 AA.
 AC Q9DFK9;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE SLENIUM BINDING PROTEIN (FRAGMENT).
 OS Gallinula chloropus (Shortjaw mudsucker).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Perciformes; Gobioidae;
 OC Gobiidae; Gobiichthys.
 OX NCBI_TaxID=79683;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE=21117151; PubMed=11172064;
 RA Gracey A.Y., Troll J.V., Somero G.N.;
 RT "Hypoxia-induced gene expression profiling in the euryoxic fish
 Gallinula chloropus mirabilis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:1993-1998(2001).
 DR EMBL: AF266226; AAG13346.1; -
 FT NON_TER 1
 SQ SEQUENCE 68 AA; 7303 MW; D6A32C267A84106A CRC64;

Query Match 3.2%; Score 15; DB 13; Length 68;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 452 GPALAHLELYPGDC 466
 |||||
 Db 48 GPALAHLELYPGDC 62

RESULT 9
 O23265 ID O23265 PRELIMINARY; PRT; 478 AA.
 AC O23265;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE HYPOPHYSAL 53.3 KDA PROTEIN (SLENIUM-BINDING PROTEIN LIKE).
 GN D13061C OR AT4G14040.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,
 RA Kreis M., Kavanagh T., Entlian K.D., Rieger W., James R.,
 RA Puidomenech P., Hatzopoulos P., Obermayer B., Duesterhoft A.,
 RA Jones J., Palme K., Ansoorge W., Delseny M., Bancroft I., Mewes H.W.,
 RA Schueller C., Chaltwats N.

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [12]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: 297335; CAB46000.1; -.
DR EMBL: AL161537; CAB78446.1; -.
DR InterPro: IPR002885; PPR.
DR InterPro: IPR005054; RRM.
DR Pfam: PF01535; PPR: 6.
DR PROSITE: PS00303; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 478 AA; 53275 MW; FCB942FFB9FD82B CRC64;

Query Match 3.2%; Score 15; DB 10; Length 478;
Best Local Similarity 100.0%; Pred. No. 9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 GPOMTOLSDGKRLY 400
Db 392 GPOMTOLSDGKRLY 406
|||||

RESULT 10
O9ASS5 PRELIMINARY: PRT: 487 AA.
AC O9ASS5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE ATG14040/DL3061C.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kamita A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,
RA Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis cDNA clones."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF367311; AAK32898.1; -.
DR InterPro: IPR000345; CytoC_heme_bind.
DR InterPro: IPR000504; RRM.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS00303; RRM_RNP_1; UNKNOWN_1.
SQ SEQUENCE 487 AA; 53965 MW; DDBB49DB527F1DBF CRC64;

Query Match 3.2%; Score 15; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 GPOMTOLSDGKRLY 400
Db 401 GPOMTOLSDGKRLY 415
|||||

RESULT 11
O93WNO PRELIMINARY: PRT: 487 AA.
AC O93WNO;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PURATIVE SILENIUM-BINDING PROTEIN.
GN ATG14040.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
RA Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
RA Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamita A.,
RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene ATG14040."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamita A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Full length cDNA of gene ATG14040 (GI:7268108)."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY040046; AAK64104.1; -.
DR EMBL: AF370130; AAK3945.1; -.
SQ SEQUENCE 487 AA; 53937 MW; 0567C4138EF78A48 CRC64;

Query Match 3.2%; Score 15; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 9.2e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 386 GPOMTOLSDGKRLY 400
Db 401 GPOMTOLSDGKRLY 415
|||||

RESULT 12
O9AVAG PRELIMINARY: PRT: 457 AA.
AC O9AVAG;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE PURATIVE SILENIUM-BINDING PROTEIN.
GN OS SBP.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; liliopsids; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. AKITAKOMACHI; TISSUE=LEAF;
RA Sawada K., Iwata M.;
RT "putative selenium binding protein - Oryza sativa."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB059401; BAB40923.1; -.
DR InterPro: IPR000345; CytoC_heme_bind.
DR InterPro: IPR001202; WW.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS01159; WW_DOMAIN_1; UNKNOWN_1.
SQ SEQUENCE 457 AA; 50957 MW; 963002CC6614D220 CRC64;

Query Match 3.0%; Score 14; DB 10; Length 457;
Best Local Similarity 100.0%; Pred. No. 9.8e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 385 GGPOMIQLSLDGKR 398
 |||||||||
 DB 369 GGPOMIQLSLDGKR 382

RESULT 13

Q93YH7 093YH7 PRELIMINARY; PRT; 478 AA.
 AC 093YH7
 DT 01-DEC-2001 (TREMBLERL. 19, Created)
 DT 01-DEC-2001 (TREMBLERL. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLERL. 19, Last annotation update)
 DE SELENIUM BINDING PROTEIN (FRAGMENT).
 GN SBP.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OX NCBI_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Flemetakis E.;
 RT "A novel Lotus japonicus nodulin gene is highly conserved among plants
 RT and animals, and encodes a homologue to the mammalian selenium binding
 RT protein.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ401230; CAC67472.1; -.
 FT NON_TER 1 1
 SQ SEQUENCE 478 AA; 53064 MW; 2A17803144985DCC CRC64;

Query Match 2.8%; Score 13; DB 10; Length 478;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 386 GGPOMIQLSLDGKR 398
 |||||||||
 DB 392 GGPOMIQLSLDGKR 404

RESULT 14
 Q42178 042178 PRELIMINARY; PRT; 146 AA.
 AC 042178;
 DT 01-NOV-1996 (TREMBLERL. 01, Created)
 DT 01-NOV-1996 (TREMBLERL. 01, Last sequence update)
 DT 01-DEC-2001 (TREMBLERL. 19, Last annotation update)
 DE SELENIUM BINDING PROTEIN (FRAGMENT).
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=DRY SEEDS OF A THALIANA ECOTYPE COLUMBIA;
 RA Raynal M., Grellier F., Laudie M., Meyer Y., Cooke R., Delseny M.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z27257; CAA81769.1; -.
 DR InterPro: IPR000345; CYTC_heme_bind.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 FT NON_TER 146 146
 SQ SEQUENCE 146 AA; 15446 MW; 5EC67EBC517808D5 CRC64;

Query Match 2.3%; Score 11; DB 10; Length 146;
 Best Local Similarity 100.0%; Pred. No. 0.0049;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 PDYLATVDVDP 51
 |||||||||
 DB 58 PDYLATVDVDP 68

RESULT 15
 Q93VAL1 093VAL1 PRELIMINARY; PRT; 487 AA.
 AC 093VAL1
 DT 01-DEC-2001 (TREMBLERL. 19, Created)
 DT 01-DEC-2001 (TREMBLERL. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLERL. 19, Last annotation update)
 DE SELENIUM BINDING PROTEIN.
 GN SBP.
 OS Lotus japonicus.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae; Lotus.
 OX NCBI_TaxID=34305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Flemetakis E.;
 RT "A novel Lotus japonicus nodulin gene is highly conserved among plants
 RT and animals, and encodes a homologue to the mammalian selenium binding
 RT protein.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ401227; CAC67492.1; -.
 DR EMBL: AJ401226; CAC67491.1; -.
 SQ SEQUENCE 487 AA; 53652 MW; DEDAC31A908CC12E CRC64;

Query Match 2.3%; Score 11; DB 10; Length 487;
 Best Local Similarity 100.0%; Pred. No. 0.015;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 PDYLATVDVDP 51
 |||||||||
 DB 55 PDYLATVDVDP 65

Search completed: August 29, 2002, 16:56:10
 Job time: 244 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:42:46 ; Search time 33.15 Seconds

(without alignments)
1581.503 Million cell updates/sec

Title: US-09-841-758-1

Perfect score: 2550
Sequence: 1 MATKCGNCGPGYSTPLEAMK.....PALAHLRLRYPGDCCSSDIWI 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_032802:*
1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2550	100.0	472	19 AAW37940	Amino acid sequence
2	2550	100.0	472	21 AAY68328	Amotrophic latera
3	2539	99.6	499	21 AAB57139	Human prostate can
4	1740.5	68.3	582	22 AAU30396	Novel human secret
5	1587.5	62.3	476	21 AAG52477	Arabidopsis thalia
6	1587.5	62.3	480	21 AAG52476	Arabidopsis thalia
7	1571.5	61.6	490	21 AAG36718	Arabidopsis thalia
8	1562.5	61.3	488	21 AAG17315	Arabidopsis thalia
9	1562.5	61.3	503	21 AAG17314	Arabidopsis thalia
10	1558.5	60.3	455	21 AAG52478	Arabidopsis thalia
11	1523.5	59.7	455	21 AAG36719	Arabidopsis thalia

12	1514.5	59.4	455	21 AAG17316	Arabidopsis thalia
13	1444.5	56.6	469	21 AAG19252	Arabidopsis thalia
14	1444.5	56.6	502	21 AAG19251	Arabidopsis thalia
15	1359.5	53.3	409	21 AAG36720	Arabidopsis thalia
16	1339.5	52.5	409	21 AAG19253	Arabidopsis thalia
17	1247.5	48.9	468	22 ABB71631	Drosophila melanog
18	1178	46.2	235	22 AAB95579	Human protein sequ
19	1079.5	42.3	327	21 AAG38104	Arabidopsis thalia
20	1073.5	42.1	328	21 AAG16909	Arabidopsis thalia
21	916.5	35.9	275	21 AAG38105	Arabidopsis thalia
22	913.5	35.8	275	21 AAG25098	Arabidopsis thalia
23	905.5	35.5	275	21 AAG16910	Arabidopsis thalia
24	807	31.6	236	21 AAG25099	Arabidopsis thalia
25	799	31.3	236	21 AAG16911	Arabidopsis thalia
26	713	28.0	199	21 AAG25100	Arabidopsis thalia
27	713	28.0	199	21 AAG38106	Arabidopsis thalia
28	686	26.9	188	21 AAG19617	Arabidopsis thalia
29	616	24.2	167	21 AAG19618	Arabidopsis thalia
30	495	19.4	138	21 AAB54427	Human pancreatic c
31	420	16.5	93	22 AAO08400	Human polypeptide
32	176	6.9	92	22 AAU32310	Novel human secret
33	146.5	5.7	59	22 AAU30395	Novel human secret
34	126	4.9	21	21 AAY68338	Amotrophic latera
35	121	4.7	20	21 AAY68332	Amotrophic latera
36	116	4.5	834	22 AAG25296	Novel human diagno
37	108.5	4.3	3470	22 AAG25297	Erythrovirus V9 VP
38	106.5	4.2	554	20 AAY32320	Erythrovirus V9 VP
39	106.5	4.2	781	20 AAY32327	Erythrovirus V9 VP
40	102.5	4.0	554	16 AAW08987	Human parvovirus V
41	102.5	4.0	657	21 AAG30622	Arabidopsis thalia
42	102.5	4.0	712	21 AAG30621	Arabidopsis thalia
43	102.5	4.0	720	21 AAG30620	Arabidopsis thalia
44	102.5	4.0	756	21 AAY71231	Capsid protein enc
45	102.5	4.0	781	16 AAW08986	Human parvovirus V

ALIGNMENTS

RESULT 1	AAW37940	standard; Protein; 472 AA.
ID	AAW37940	
XX	AAW37940	
AC	AAW37940	
XX		
DT	10-AUG-1998	(first entry)
XX		
DE		Amino acid sequence of the human selenium-binding protein.
KW		Human selenium-binding protein; HSEBP; recombinant HSEBP; liver damage;
KW		kidney damage; lung damage; antibody; cancer.
XX		
OS		Homo sapiens.
XX		
PN	US5759812-A.	
XX		
PD	02-JUN-1998.	
XX		
FE	15-NOV-1996;	96US-0749903.
XX		
PR	15-NOV-1996;	96US-0749903.
XX		
PA	(INCY-) INCYTE PHARM INC.	
XX		
PI	Bandman O, Hawkins PR;	
XX		
DR	WPI: 1998-332133/29.	
XX		
XX	N-PSDB: AAV29245.	
PT	DNA encoding human selenium-binding protein - useful for producing	
PT	recombinant protein	
XX		
PS	Claim 1; Columns 33-36; 35pp; English.	

XX This is the amino acid sequence of the novel human selenium-binding
CC protein (HSEBP). In the method of the invention, recombinant HSEBP
CC is produced which can be used to treat liver, kidney or lung damage
CC caused by chemical agents, or to produce antibodies. The vector can
CC be used to treat liver, kidney or lung damage caused by chemical
CC agents or to treat cancer.

XX Sequence 472 AA;

Query Match 100.0%; Score 2550; DB 19; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.3e-249;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGPGYSTPLEAMKGPREEIVLPICIRNTGTEAPDYLATVDKSPQYCVI 60
DB 1 matkcgngpgystpleamkgspreelivlpicirntgteapdylatvdqpspycqv1 60
QY 61 HRLPMPNLKDELHSGMWTSCSFGSDSTKSRKLVLPSSLISRIYVVDGSEPRAPKLHK 120
DB 61 hrlpmpnlkdelhsgmntscsfgsdstksrklvlpssliisriyvvdgseprapklhk 120
QY 121 VIEPKDIAKCELAFLHSHCLASGEVMISSLDGVKNGKGFEVLLDGEFEVKGTEWERP 180
DB 121 viepkdihakcelafhshclasevwmisldgvkngkgfsvllldgefevkgtewerp 180
QY 181 GGAAPLGIDFWYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYSHLYVMDQHEIYQ 240
DB 181 ggaaplgidfwyqprhnwmistemaapnvlrdgfnpadveaglyshlyvmdqheiyq 240
QY 241 TSLKDGILPIELIRFLHNPDAQGFVGCALSTIORFYKNEGTSVEKVIQVPPKKYKG 300
DB 241 tslkdgilpielirflhnpdaagfvgcalstiorfyknegtsvekvivppkkkyg 300
QY 301 WLPPMPGLITDILLSLDRFLYFSNWLHGDROYDISDPPRRLTGOFLGSIYKGGP 360
DB 301 wlppmpglitdillslldrflfysnwlhgdroydisdpprrltgoflgsiykggp 360
QY 361 VOVLDEDELSKSPERLVYKGRVAGPQMIQSLDGKRLYITTSLSXAMDKQFYFDLIRE 420
DB 361 vovldeelskspervlvkgrvagpomiqlsdgkrllyittslysawdkqfyfdlire 420
QY 421 GSVMLQVVDVTVKGGKLNPNFLVDFGKEPLPALAHELRYGDCSSDIWI 472
DB 421 gsvmlqvvdvkvkgklnpnflvdfgkeplpalahelrypgdcssdiwi1 472

RESULT 2

AAV68328 standard: Protein: 472 AA.

AAV68328;

DT 17-APR-2000 (first entry)

XX Amyotrophic lateral sclerosis related p53 protein.

XX Amyotrophic lateral sclerosis; p53; ALS; diagnosis; familial ALS.

OS Homo sapiens.

PN JP2000000095-A.

PD 07-JAN-2000.

PF 15-JUN-1998; 98JP-0167259.

PR 15-JUN-1998; 98JP-0167259.

PA (SAKO/) SAKODA S

XX (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.

DR WPI: 2000-129429/12.

DR N-PSDB: AAZ57589.

PT New DNA encoding the p53 protein - useful in the diagnosis of

PT amyotrophic lateral sclerosis

PS Claim 1; Page 6-8; 11pp: Japanese.

CC The present sequence represents the protein p53 which is related to
CC amyotrophic lateral sclerosis (ALS). The protein p53 and its gene
CC can be used for the diagnosis of ALS.

SQ Sequence 472 AA;

Query Match 100.0%; Score 2550; DB 21; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.3e-249;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGPGYSTPLEAMKGPREEIVLPICIRNTGTEAPDYLATVDKSPQYCVI 60
DB 1 matkcgngpgystpleamkgspreelivlpicirntgteapdylatvdqpspycqv1 60
QY 61 HRLPMPNLKDELHSGMWTSCSFGSDSTKSRKLVLPSSLISRIYVVDGSEPRAPKLHK 120
DB 61 hrlpmpnlkdelhsgmntscsfgsdstksrklvlpssliisriyvvdgseprapklhk 120
QY 121 VIEPKDIAKCELAFLHSHCLASGEVMISSLDGVKNGKGFEVLLDGEFEVKGTEWERP 180
DB 121 viepkdihakcelafhshclasevwmisldgvkngkgfsvllldgefevkgtewerp 180
QY 181 GGAAPLGIDFWYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYSHLYVMDQHEIYQ 240
DB 181 ggaaplgidfwyqprhnwmistemaapnvlrdgfnpadveaglyshlyvmdqheiyq 240
QY 241 TSLKDGILPIELIRFLHNPDAQGFVGCALSTIORFYKNEGTSVEKVIQVPPKKYKG 300
DB 241 tslkdgilpielirflhnpdaagfvgcalstiorfyknegtsvekvivppkkkyg 300
QY 301 WLPPMPGLITDILLSLDRFLYFSNWLHGDROYDISDPPRRLTGOFLGSIYKGGP 360
DB 301 wlppmpglitdillslldrflfysnwlhgdroydisdpprrltgoflgsiykggp 360
QY 361 VOVLDEDELSKSPERLVYKGRVAGPQMIQSLDGKRLYITTSLSXAMDKQFYFDLIRE 420
DB 361 vovldeelskspervlvkgrvagpomiqlsdgkrllyittslysawdkqfyfdlire 420
QY 421 GSVMLQVVDVTVKGGKLNPNFLVDFGKEPLPALAHELRYGDCSSDIWI 472
DB 421 gsvmlqvvdvkvkgklnpnflvdfgkeplpalahelrypgdcssdiwi1 472

RESULT 3

AAB57139 standard: Protein: 499 AA.

AC AAB57139;

DT 13-MAR-2001 (first entry)

DE Human prostate cancer antigen protein sequence SEQ ID NO:1717.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

XX neuroprotective; cytosolic; cardioprotective; immunomodulatory; muscular;

XX vulnerrary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;

XX antibacterial; gene therapy; neural; immune; reproductive; renal;

XX gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

XX wound; infectious disease.

OS Homo sapiens.

XX WO200055174-A1.

```

PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US05988.
XX
PR 12-MAR-1999; 990US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX (ROSE/) ROSEN C A.
PI Rosen CA, Ruben SM;
XX
XX WPI: 2000-587513/55.
DR N-PSDB: AAF16342.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer.
XX
XX Claim 11; Page 2190-2192; 2338pp; English.
XX
XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnerrary, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF15506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
XX
XX Sequence 499 AA;
SQ

```

Query Match 99.6%; Score 2539; DB 21; Length 499;
 Best Local Similarity 99.6%; Pred. No. 1.9e-248;
 Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MATKCGNGPGYSTPLEAMKGPREEIYVLPCTYRNTGTETAPDYLATVVDVDFPKSPQYCOV 60
  |||||
DB 28 matkcgncgpgystpleamkgpreeivlylpctyrntgtetapdylatvdvdpkspqycv 87
  |||||
QY 61 HRLPMNLKDELHSHGWNMTSCSGDSTKSRKTVLPDLSSRIYVVDVSGSEPPAPKLUK 120
  |||||
DB 88 hrlpmlkdelhshgwnmtscsgdstkstklvlpdlssriyvvdvsgseppaprkluh 147
  |||||
QY 121 VIEPKDIHAKCELAFLHTSHCLASGEWISLGDVKGNGKGFVLLDGEPEVGTWERP 180
  |||||
DB 148 viepkdihakceiaflhtshclasegwmisslgdvkgngyggvylldgetevgtwerp 207
  |||||
QY 181 GGAAPLIGYDFWYOPRANHMTSTEWAPNVLDGFNPADVEALGSHLYWDMORHEIYQ 240
  |||||
DB 208 ggaaplgydfwyoprnhmstewapnvlrdgfnpadveaglygshlywdmqrheivq 267
  |||||
QY 241 TLSKIDGLIPLERFLNPNDAAGFVGCALSTIQRFYKNEGTSVEKYIQVPPKRVKG 300
  |||||
DB 268 tlskldgliplrflnnpdaagfvgcalstlqrfyknegtsvekyiqvppkrvkg 327
  |||||
QY 301 WLDEMGGLTDDILSLDDRFLYFSNMVHGDRLROYDISDPPRPGLTQDLFGSGITVKG 360
  |||||
DB 328 wldepmpglitdillslddrflfysnmvhdgdlroydisdpprpilvqqlfsgsitvkgp 387
  |||||
QY 361 VQVLEDEELASQPEPLVYVKGKRVAGPQMTQLSLDGKRLYITTSYLSAMPKQFPPDLIRE 420
  |||||
DB 388 vqvledeelsqpeplvvkkgkrvagpqrmtqlslldgkrllyttsylsawmdqfpdlire 447
  |||||
QY 421 GSWALQVDDVTVKGGLKLNPNFLVDGKPEPLGPAHLAHLRTPGDGSSDIWT 472
  |||||
DB 448 gswalqvddvtvkgglklnpnflvdgkpeplgpalahlxrtpgdcssdiwt 499
  |||||

```

```

RESULT 4
AAU30396
ID AAU30396 standard; Protein: 582 AA.
XX
XX AAU30396;
AC
XX
XX 18-DEC-2001 (first entry)
DT
XX
XX Novel human secreted protein #887.
DE
XX
XX Human: vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200179449-A2.
PN
XX
XX 25-OCT-2001.
PD
XX
XX 16-APR-2001; 2001MO-US08656.
PF
XX
XX 18-APR-2000; 2000US-0552929.
PR
XX
XX 26-JAN-2001; 2001US-0770160.
PA
XX (HYSE-) HYSEQ INC.
PI
XX
XX Tang YT, Liu C, Drmanac RT;
PT
XX WPI: 2001-611725/70.
DR
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy.
XX
XX Claim 20; Page 288; 765pp; English.
XX
XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukaemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.
XX
XX Sequence 582 AA;
SQ

```

Query Match 68.3%; Score 1740.5; DB 22; Length 582;
 Best Local Similarity 73.5%; Pred. No. 2.6e-167;
 Matches 367; Conservative 21; Mismatches 78; Indels 33; Gaps 20;

```

QY 2 ATKCGNGPGYSTPLEAMKGPREEIYVLPCTYRNT-GTEAPDY-LATVVDVDFPKSPQYCOV 59
  |||||
DB 37 atkcgngpgystpleamkgpreeivlylpctyrntgteapdylatvvdvdpkspqycv 96
  |||||
QY 60 IHRPMPNLKDELHSHGWNMTSCSGDSTKSR-TKVLPSL-TISSRIYVVDV-GSEPPAP 116
  |||||
DB 97 ihrlpmlkdelhshgwnmtscsgdstkstklvlpdlssriyvvdvsgseppap 156
  |||||
QY 117 VIEPKDIHAKCELAFLHTSHCLASGEWISLGDVKGNGKGFVLLDGEPEVKG 175
  |||||

```

Db 157 klagsel|sprdl|hakcnwaf|htshcl|asgevm|slgdv|kngk|ggyf|ll|dget|evk|g 216
QY 176 TWERPGGAAPLCYDFWYQPRHNVMISTEWAPNVLDGFNPADVEAGLYGSHLYVWDMQR 235
Db 217 twerpggaap|lydf|wyqpr|hnvm|istewa|pnv|ldgf|npad|veag|lygsh|lyvw|dmqr 276
QY 236 HEIVQTL|SLK|DGL|PLE|EIF|LN|PD-AAQGFV|GAL|SS-T|Q|RF-Y|KNE|GCT-W|SE|EKY|I 291
Db 277 |e|iv|q|t|s|l|k|d|g|l|p|e|t|f|f|l|n|p|d|c|p|k|a|f|v|g|c|a|l|q|a|p|n|l|q|f|l|q|f|r|t|g|l|f|s|g|r|x|f 336
QY 292 QVPPKVKGMWLLPEMPGGLTDTLLS-LDDRP|LYFS|NWL-HGDL|ROYD|DISD|P|R|L|T|G|L 349
Db 337 gvp|p|k|l|k|w|l|p|k|n|p|l|t|l|a|s|w|n|d|g|f|l|y|s|n|l|a|w|g|p|x|k|y|d|s|q|p|q|a|l|t|g|l 396
QY 350 FLGSG|LVK|GPGV|QVLE|DE|ELKS-Q|EP|LV|V|K|G|R|V|A|G|P|O|M|I|Q|L|S|D|G|R|L-|Y|T|T 403
Db 397 fl|g|s|t|v|k|e|g|p|q|v|l|e|d|e|l|s|p|e|l|v|v|k|g|t|v|g|e|g|m|q|l|s|d|g|k|r|l|n|h|g|r|c|s 456
QY 404 SLYSAMD|KQ|FY|D-|L|I|REG|SV|MLQ|VDV|V|K|G-L|K|N|P|N|F-L|V|D|E|G|K-E|P|L|G|P 453
Db 457 a|l-|q|g|s|f|t|p|x|s|g|e|l|l|v|n|a|g|v-|e|w|x|d|n|s|k|g|a|x|k|l|n|p|q|l|s|v|d|f|g|e|g|a|p|l|p|k 511
QY 454 ALAHE|L|R|Y|P|G|G|C|S|D|I|W|I 472
Db 512 p|l|p|h|x|a|y|p|g|d|e|s|d|l|w|l 530

RESULT 5

AAG52477 standard: Protein; 476 AA.

AAG52477;

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 66711.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.

OS Arabidopsis thaliana.

XX
PN EP1033405-A2.XX
PD 06-SEP-2000.XX
PF 25-FEB-2000; 2000EP-0301439.XX
PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130444.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131444.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 04-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 14-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0138110.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140033.
PR 24-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.

PR	27-JUL-1999	9905-014557.6
PR	27-JUL-1999	9905-014551.3
PR	27-JUL-1999	9905-014501.8
PR	28-JUL-1999	9905-014551.9
PR	02-AUG-1999	9905-014638.6
PR	02-AUG-1999	9905-014638.8
PR	02-AUG-1999	9905-014658.9
PR	03-AUG-1999	9905-014703.8
PR	04-AUG-1999	9905-014720.4
PR	04-AUG-1999	9905-014730.2
PR	05-AUG-1999	9905-014719.2
PR	05-AUG-1999	9905-014726.0
PR	06-AUG-1999	9905-014776.3
PR	09-AUG-1999	9905-014741.6
PR	09-AUG-1999	9905-014793.5
PR	10-AUG-1999	9905-014817.1
PR	12-AUG-1999	9905-014831.9
PR	13-AUG-1999	9905-014854.1
PR	13-AUG-1999	9905-014856.5
PR	16-AUG-1999	9905-014968.4
PR	17-AUG-1999	9905-014917.5
PR	18-AUG-1999	9905-014942.6
PR	20-AUG-1999	9905-014972.3
PR	20-AUG-1999	9905-014972.3
PR	20-AUG-1999	9905-014992.0
PR	23-AUG-1999	9905-014993.0
PR	23-AUG-1999	9905-014993.0
PR	26-AUG-1999	9905-015066.6
PR	27-AUG-1999	9905-015008.4
PR	27-AUG-1999	9905-015100.6
PR	27-AUG-1999	9905-015106.6
PR	30-AUG-1999	9905-015130.3
PR	31-AUG-1999	9905-015130.3
PR	01-SEP-1999	9905-015133.0
PR	07-SEP-1999	9905-015207.0
PR	10-SEP-1999	9905-015306.3
PR	13-SEP-1999	9905-015375.8
PR	15-SEP-1999	9905-015401.8
PR	16-SEP-1999	9905-015403.9
PR	20-SEP-1999	9905-015477.9
PR	22-SEP-1999	9905-015513.9
PR	24-SEP-1999	9905-015535.8
PR	28-SEP-1999	9905-015645.8
PR	28-SEP-1999	9905-015656.9
PR	04-OCT-1999	9905-015717.3
PR	05-OCT-1999	9905-015751.5
PR	06-OCT-1999	9905-015786.5
PR	07-OCT-1999	9905-015802.9
PR	08-OCT-1999	9905-015823.2
PR	12-OCT-1999	9905-015836.9
PR	13-OCT-1999	9905-015929.3
PR	13-OCT-1999	9905-015929.4
PR	13-OCT-1999	9905-015929.9
PR	14-OCT-1999	9905-015933.1
PR	14-OCT-1999	9905-015963.7
PR	14-OCT-1999	9905-015963.8
PR	18-OCT-1999	9905-015980.4
PR	21-OCT-1999	9905-016007.1
PR	21-OCT-1999	9905-016017.9
PR	21-OCT-1999	9905-016070.8
PR	21-OCT-1999	9905-016076.6
PR	21-OCT-1999	9905-016081.5
PR	22-OCT-1999	9905-016098.0
PR	22-OCT-1999	9905-016098.1
PR	22-OCT-1999	9905-016098.9

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match	62.3%	Score 1587.5	DB 21	length 476
Best Local Similarity	61.3%	Pred. No. 6.3e-152		
Matches 288	Conservative	63	Mismatches 116	Indels 3
				Gaps 3

QY 5 CGNCGGYSFLEAMKGPREEIVYLPCIRANTGTAEADPLATVVDVDPSPQYQGVHILP 64
| | | | | : | | | | |
Db 8 ccksgpyatpllamsprekllyaaalytqgkpkdylatvdvpepsstlyssvinlilp 67

[illegible][illegible]

OY
184 APLGVEWYQGRHNWMISTEAAAPNVLRGEFNPAEDVGAALYSYLXVDMDORHEIVCLLS 243
| | | | | : | | | | | : | | | | | :
Db
187 pLYgydfwqprhkumstswgapaaftkfgldkdvsdglgyskhnhvswpgelkqlld 246

[illegible]

302 1pmpm1itcdlslslddrflgyscmwhgdrtymledprkvtlqglmgvdlvqkgsavl 366

```

367  algeegafqdvptkqgrlrgspqmfgslgdkrrilyvnslstswdqfypelveks  426

```

Db 427 hmliqivdtkgslslnpnlffvdfstegpagslahemrypggdcstslw 476

RESULT 6
AAG52476
ID AAG52476 standard; Protein; 480 AA.

AC	AA652476;
XX	
DT	18-OCT-2000 (first entry)
XX	

xx Arabidopsis thaliana protein fragment SEQ ID NO: 66710.
xx
xx Protein identification: signal transduction pathway; metabolic pathway;
xx hybridization assay; open pit; maternal gene; defense; control;
xx growth; development; cell cycle; cell division; cell death; cell survival;
xx

KW termination sequence.
XX
OS *Arabidopsis thaliana*.
XY

PN EPI033405-A2.
XX
XX
PD 06-SEP-2000.
XX

AA	9905-0121825.
PR	25-FEB-1999:
PR	9905-0123180.
PR	05-MAR-1999:
PR	9905-0123548.
PR	09-MAR-1999:

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR	18-JUN-1999	99S05-0139463
PR	18-JUN-1999	99S05-0139463
PR	18-JUN-1999	99S05-0139763
PR	21-JUN-1999	99S05-0139617
PR	22-JUN-1999	99S05-0139699
PR	23-JUN-1999	99S05-0140354
PR	23-JUN-1999	99S05-0140354
PR	24-JUN-1999	99S05-0140695
PR	24-JUN-1999	99S05-0140623
PR	29-JUN-1999	99S05-0140991
PR	30-JUN-1999	99S05-0141887
PR	30-JUN-1999	99S05-0141842
PR	01-JUL-1999	99S05-0142154
PR	02-JUL-1999	99S05-0142051
PR	06-JUL-1999	99S05-0142390
PR	08-JUL-1999	99S05-0142803
PR	09-JUL-1999	99S05-0142820
PR	12-JUL-1999	99S05-0142877
PR	13-JUL-1999	99S05-0143542
PR	14-JUL-1999	99S05-0143624
PR	15-JUL-1999	99S05-0144005
PR	16-JUL-1999	99S05-0144085
PR	16-JUL-1999	99S05-0144085
PR	19-JUL-1999	99S05-0144325
PR	19-JUL-1999	99S05-0144331
PR	19-JUL-1999	99S05-0144331
PR	19-JUL-1999	99S05-0144331
PR	19-JUL-1999	99S05-0144334
PR	19-JUL-1999	99S05-0144335
PR	20-JUL-1999	99S05-0144335
PR	20-JUL-1999	99S05-0144332
PR	20-JUL-1999	99S05-0144384
PR	21-JUL-1999	99S05-0144614
PR	21-JUL-1999	99S05-0145086
PR	21-JUL-1999	99S05-0145086
PR	21-JUL-1999	99S05-0145087
PR	22-JUL-1999	99S05-0145087
PR	22-JUL-1999	99S05-0145089
PR	22-JUL-1999	99S05-0145192
PR	23-JUL-1999	99S05-0145145
PR	23-JUL-1999	99S05-0145214
PR	23-JUL-1999	99S05-0145224
PR	26-JUL-1999	99S05-0145376
PR	27-JUL-1999	99S05-0145513
PR	27-JUL-1999	99S05-0145518
PR	27-JUL-1999	99S05-0145519
PR	28-JUL-1999	99S05-0145511
PR	02-AUG-1999	99S05-0146386
PR	02-AUG-1999	99S05-0146388
PR	02-AUG-1999	99S05-0146388
PR	06-AUG-1999	99S05-0147416
PR	09-AUG-1999	99S05-0147335
PR	10-AUG-1999	99S05-0148171
PR	11-AUG-1999	99S05-0148319
PR	12-AUG-1999	99S05-0148341
PR	13-AUG-1999	99S05-0148345
PR	13-AUG-1999	99S05-0148684
PR	16-AUG-1999	99S05-0149368
PR	17-AUG-1999	99S05-0149175
PR	18-AUG-1999	99S05-0149426
PR	20-AUG-1999	99S05-0149222
PR	20-AUG-1999	99S05-0149723
PR	20-AUG-1999	99S05-0149929
PR	23-AUG-1999	99S05-0149902
PR	25-AUG-1999	99S05-0150066
PR	25-AUG-1999	99S05-0150566

PR	28-AUG-1999	9905-0150864
PR	27-AUG-1999	9905-0150865
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151067
PR	30-AUG-1999	9905-0151103
PR	31-AUG-1999	9905-0151130
PR	01-SEP-1999	9905-0151348
PR	07-SEP-1999	9905-0152633
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154039
PR	20-SEP-1999	9905-0154779
PR	22-SEP-1999	9905-0155133
PR	23-SEP-1999	9905-0155486
PR	24-SEP-1999	9905-0155659
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	06-OCT-1999	9905-0157865
PR	07-OCT-1999	9905-0158029
PR	08-OCT-1999	9905-0158322
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	18-OCT-1999	9905-0159638
PR	18-OCT-1999	9905-0159684
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160747
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160774
PR	21-OCT-1999	9905-0160810
PR	21-OCT-1999	9905-0160815
PR	22-OCT-1999	9905-0160980
PR	22-OCT-1999	9905-0160981
PR	22-OCT-1999	9905-0160989
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161406
PR	25-OCT-1999	9905-0161407
PR	26-OCT-1999	9905-0161359
PR	26-OCT-1999	9905-0161360
PR	26-OCT-1999	9905-0161361
PR	28-OCT-1999	9905-0161920
PR	28-OCT-1999	9905-0161932
PR	28-OCT-1999	9905-0161942
PR	29-OCT-1999	9905-0162143

Query Match 61.6%; Score 1571.5; DB 21; Length 490;

Best Local Similarity 61.18; Pred. No. 2.8e-150;
Matches 287; Conservative 65; Mismatches 115; Indels 3; Gaps 3;

[illegible]

04-AUG-1999;	99US-0147302.
PR	
05-AUG-1999;	99US-0147192.
PR	
05-AUG-1999;	99US-0147260.
PR	
06-AUG-1999;	99US-0147303.
PR	
06-AUG-1999;	99US-0147416.
PR	
09-AUG-1999;	99US-0147493.
PR	
09-AUG-1999;	99US-0147935.
PR	
10-AUG-1999;	99US-0148171.
PR	
11-AUG-1999;	99US-0148319.
PR	
12-AUG-1999;	99US-0148341.
PR	
13-AUG-1999;	99US-0148565.
PR	
13-AUG-1999;	99US-0148684.
PR	
16-AUG-1999;	99US-0149368.
PR	
17-AUG-1999;	99US-0149175.
PR	
18-AUG-1999;	99US-0149426.
PR	
20-AUG-1999;	99US-0149722.
PR	
20-AUG-1999;	99US-0149723.
PR	
20-AUG-1999;	99US-0149929.
PR	
23-AUG-1999;	99US-0149902.
PR	
23-AUG-1999;	99US-0149930.
PR	
25-AUG-1999;	99US-0150566.
PR	
26-AUG-1999;	99US-0150884.
PR	
27-AUG-1999;	99US-0151065.
PR	
27-AUG-1999;	99US-0151066.
PR	
27-AUG-1999;	99US-0151080.
PR	
30-AUG-1999;	99US-0151303.
PR	
31-AUG-1999;	99US-0151438.
PR	
01-SEP-1999;	99US-0151930.
PR	
07-SEP-1999;	99US-0152363.
PR	
10-SEP-1999;	99US-0153070.
PR	
13-SEP-1999;	99US-0153758.
PR	
15-SEP-1999;	99US-0154018.
PR	
16-SEP-1999;	99US-0154039.
PR	
20-SEP-1999;	99US-0154779.
PR	
22-SEP-1999;	99US-0155139.
PR	
23-SEP-1999;	99US-0155486.
PR	
24-SEP-1999;	99US-0155659.
PR	
28-SEP-1999;	99US-0156458.
PR	
29-SEP-1999;	99US-0156596.
PR	
04-OCT-1999;	99US-0157117.
PR	
05-OCT-1999;	99US-0157753.
PR	
06-OCT-1999;	99US-0157865.
PR	
07-OCT-1999;	99US-0158029.
PR	
08-OCT-1999;	99US-0158232.
PR	
12-OCT-1999;	99US-0158369.
PR	
13-OCT-1999;	99US-0159293.
PR	
13-OCT-1999;	99US-0159294.
PR	
13-OCT-1999;	99US-0159295.
PR	
14-OCT-1999;	99US-0159329.
PR	
14-OCT-1999;	99US-0159330.
PR	
14-OCT-1999;	99US-0159331.
PR	
14-OCT-1999;	99US-0159637.
PR	
14-OCT-1999;	99US-0159638.
PR	
18-OCT-1999;	99US-0159584.
PR	
21-OCT-1999;	99US-0160741.
PR	
21-OCT-1999;	99US-0160767.
PR	
21-OCT-1999;	99US-0160778.
PR	
21-OCT-1999;	99US-0160770.
PR	
21-OCT-1999;	99US-0160814.
PR	
21-OCT-1999;	99US-0160815.
PR	
22-OCT-1999;	99US-0160980.
PR	
22-OCT-1999;	99US-0160981.
PR	
22-OCT-1999;	99US-0160989.
PR	
22-OCT-1999;	99US-0160989.
PR	
25-OCT-1999;	99US-0161404.
PR	
25-OCT-1999;	99US-0161405.
PR	
25-OCT-1999;	99US-0161406.
PR	
26-OCT-1999;	99US-0161359.
PR	
26-OCT-1999;	99US-0161360.
PR	
26-OCT-1999;	99US-0161361.
PR	
28-OCT-1999;	99US-0161920.
PR	
28-OCT-1999;	99US-0161992.
PR	
28-OCT-1999;	99US-0161993.
PR	
29-OCT-1999;	99US-0162142.
PR	
04-AUG-1999;	99US-0147302.
PR	
05-AUG-1999;	99US-0147192.
PR	
05-AUG-1999;	99US-0147260.
PR	
06-AUG-1999;	99US-0147303.
PR	
06-AUG-1999;	99US-0147416.
PR	
09-AUG-1999;	99US-0147493.
PR	
09-AUG-1999;	99US-0147935.
PR	
10-AUG-1999;	99US-0148171.
PR	
11-AUG-1999;	99US-0148319.
PR	
12-AUG-1999;	99US-0148341.
PR	
13-AUG-1999;	99US-0148565.
PR	
13-AUG-1999;</	

PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140635.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142977.
 PR 12-JUL-1999; 99US-0143542.
 PR 13-JUL-1999; 99US-0143624.
 PR 14-JUL-1999; 99US-0144005.
 PR 15-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151438.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159325.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.

```
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161933.
PR 29-OCT-1999; 990S-0162142.
```

Query Match Best Local Similarity 61.3%; Score 1562.5; DB 21; Length 503;

Matches 283; Conservative 66; Mismatches 116; Indels 3; Gaps 3;

```
OY 5 GCNCGPGVSTPLEAMKGPREEIVLYPCITRYNTGTEAPDYLATVDVDPKSPQYCOVTHRLP 64
DB 35 ckygpgpyvatpilaamsqpekilyltavyltgtdkpdylatvdvdpsspsvsvhlrip 94
OY 65 MNLKDELHSGWNTCCSGFGDSTSRKLYLPSLISSKIYVVDVNGSEPRARKLHKVIEP 124
DB 95 mpfvgdelhsgwmscscshgavsdrylvlpsslsqgrylaidtkemprrapslykyvdp 154
OY 125 KDIHAKCELAFHTSHCLASGEWMTSLGDVKGNGK-GFVLLDGEFTEFKTWERPQGA 183
DB 155 keiadktgtlafhtahclactgeilvscjgdeegnakngfllldsd-friknrwekpgns 213
OY 184 ABLGYDFWYQPRHNYMISTEMAPVNLRDGFNPADVEAGLYGSHLYWDMORHEIYQTL 243
DB 214 plygydfwypqrhkcmtstswgarkafskgfnlghvadglyshlhvyswpgemqkllid 273
OY 244 L-KDGLIPLEIFRFLNPDAAGFVGCALSTIQRFYKNGCWSEKVLQVPRKVKKML 302
DB 274 lqptglipleirflhnpbskdtgfvgsalssnmhlfiknsdelwshenvlsvxplkvenw 333
OY 303 LPEMGLITDILSLDDRFLYFSNMLHGDLROYDISDPQRPRLTGOLFVGSSTVKGPGVQ 362
DB 334 lpempglltdflisidrtfiyfvnwlnghdlrqnledpkrpvlteqilwvgallqkyspvk 393
OY 363 VLEDEELKSQPEPLVYKGRVAGCPOMIQLSIDGKRLLYTTSTLYSAMDQKQFPDLIREGS 422
DB 394 avgedgnftqfvevpqlkgsitrgspmqqlsldgkrllyatnslfawwdtqfpeimekgs 453
OY 423 VMLQVDVDTVKGGLKLNPLFVDEGKEPLGRLAHLKPPGDCSSDIWI 472
DB 454 hliqdvtekgqltlnpdlfvdgdepdgsiahemrypgdcsdwl 503
```

RESULT 10

AA652478 AAG52478 standard; Protein; 455 AA.

AA652478;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 66712.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX

```
OS Arabidopsis thaliana.
XX
XX EPI033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 990S-0121825.
XX 05-MAR-1999; 990S-012180.
XX 09-MAR-1999; 990S-0123548.
XX 23-MAR-1999; 990S-0125788.
XX 25-MAR-1999; 990S-0126264.
XX 29-MAR-1999; 990S-0126785.
XX 01-APR-1999; 990S-0127462.
XX 06-APR-1999; 990S-0128234.
XX 08-APR-1999; 990S-0128714.
XX 16-APR-1999; 990S-0129845.
XX 19-APR-1999; 990S-0130077.
XX 21-APR-1999; 990S-0130449.
XX 23-APR-1999; 990S-0130510.
XX 23-APR-1999; 990S-0130891.
XX 28-APR-1999; 990S-0131449.
XX 30-APR-1999; 990S-0132048.
XX 30-APR-1999; 990S-0132407.
XX 04-MAY-1999; 990S-0132484.
XX 05-MAY-1999; 990S-0132485.
XX 06-MAY-1999; 990S-0132486.
XX 06-MAY-1999; 990S-0132487.
XX 07-MAY-1999; 990S-0132863.
XX 11-MAY-1999; 990S-0134256.
XX 14-MAY-1999; 990S-0134218.
XX 14-MAY-1999; 990S-0134219.
XX 14-MAY-1999; 990S-0134221.
XX 14-MAY-1999; 990S-0134370.
XX 18-MAY-1999; 990S-0134768.
XX 19-MAY-1999; 990S-0134941.
XX 20-MAY-1999; 990S-0135124.
XX 21-MAY-1999; 990S-0135353.
XX 24-MAY-1999; 990S-0135629.
XX 25-MAY-1999; 990S-0136021.
XX 27-MAY-1999; 990S-0136392.
XX 28-MAY-1999; 990S-0136782.
XX 01-JUN-1999; 990S-0137222.
XX 03-JUN-1999; 990S-0137528.
XX 04-JUN-1999; 990S-0137502.
XX 07-JUN-1999; 990S-0137724.
XX 08-JUN-1999; 990S-0138094.
XX 10-JUN-1999; 990S-0138540.
XX 10-JUN-1999; 990S-0138847.
XX 14-JUN-1999; 990S-0139119.
XX 16-JUN-1999; 990S-0139452.
XX 16-JUN-1999; 990S-0139453.
XX 17-JUN-1999; 990S-0139492.
XX 18-JUN-1999; 990S-0139454.
XX 18-JUN-1999; 990S-0139455.
XX 18-JUN-1999; 990S-0139456.
XX 18-JUN-1999; 990S-0139457.
XX 18-JUN-1999; 990S-0139461.
XX 18-JUN-1999; 990S-0139462.
XX 18-JUN-1999; 990S-0139463.
XX 18-JUN-1999; 990S-0139750.
XX 18-JUN-1999; 990S-0139763.
XX 21-JUN-1999; 990S-0139817.
XX 22-JUN-1999; 990S-0139899.
XX 23-JUN-1999; 990S-0140353.
XX 23-JUN-1999; 990S-0140354.
XX 24-JUN-1999; 990S-0140695.
XX 28-JUN-1999; 990S-0140823.
XX 29-JUN-1999; 990S-0140991.
```

PR	04-JUN-1999;	9905-0114128;
PR	01-JUL-1999;	9905-0114184;
PR	01-JUL-1999;	9905-0142154;
PR	02-JUL-1999;	9905-0142155;
PR	06-JUL-1999;	9905-0142300;
PR	08-JUL-1999;	9905-0142803;
PR	09-JUL-1999;	9905-0142917;
PR	12-JUL-1999;	9905-0142977;
PR	13-JUL-1999;	9905-0143524;
PR	14-JUL-1999;	9905-0143624;
PR	15-JUL-1999;	9905-0144005;
PR	16-JUL-1999;	9905-0144085;
PR	16-JUL-1999;	9905-0144086;
PR	19-JUL-1999;	9905-0144325;
PR	19-JUL-1999;	9905-0144331;
PR	19-JUL-1999;	9905-0144332;
PR	19-JUL-1999;	9905-0144333;
PR	19-JUL-1999;	9905-0144334;
PR	20-JUL-1999;	9905-0144335;
PR	20-JUL-1999;	9905-0144352;
PR	20-JUL-1999;	9905-0144632;
PR	20-JUL-1999;	9905-0144684;
PR	21-JUL-1999;	9905-0144681;
PR	21-JUL-1999;	9905-0144686;
PR	21-JUL-1999;	9905-0145087;
PR	21-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145085;
PR	22-JUL-1999;	9905-0145087;
PR	22-JUL-1999;	9905-0145088;
PR	22-JUL-1999;	9905-0145089;
PR	22-JUL-1999;	9905-0145102;
PR	23-JUL-1999;	9905-0145145;
PR	23-JUL-1999;	9905-0145218;
PR	23-JUL-1999;	9905-0145224;
PR	26-JUL-1999;	9905-0145276;
PR	27-JUL-1999;	9905-01452913;
PR	27-JUL-1999;	9905-01452918;
PR	27-JUL-1999;	9905-01452919;
PR	28-JUL-1999;	9905-0145951;
PR	02-AUG-1999;	9905-0146386;
PR	02-AUG-1999;	9905-0146388;
PR	03-AUG-1999;	9905-0146389;
PR	03-AUG-1999;	9905-0146389;
PR	04-AUG-1999;	9905-0147204;
PR	04-AUG-1999;	9905-0147302;
PR	05-AUG-1999;	9905-0147192;
PR	05-AUG-1999;	9905-0147260;
PR	06-AUG-1999;	9905-0147303;
PR	06-AUG-1999;	9905-0147416;
PR	09-AUG-1999;	9905-0147493;
PR	09-AUG-1999;	9905-0147935;
PR	10-AUG-1999;	9905-0148171;
PR	11-AUG-1999;	9905-0148319;
PR	12-AUG-1999;	9905-0148341;
PR	13-AUG-1999;	9905-0148565;
PR	13-AUG-1999;	9905-0148684;
PR	16-AUG-1999;	9905-0149378;
PR	17-AUG-1999;	9905-0149131;
PR	18-AUG-1999;	9905-0149426;
PR	20-AUG-1999;	9905-0149722;
PR	20-AUG-1999;	9905-0149723;
PR	23-AUG-1999;	9905-0149929;
PR	23-AUG-1999;	9905-0149902;
PR	23-AUG-1999;	9905-0149930;
PR	25-AUG-1999;	9905-0150566;
PR	26-AUG-1999;	9905-0150884;
PR	27-AUG-1999;	9905-0151065;
PR	27-AUG-1999;	9905-0151066;
PR	27-AUG-1999;	9905-0151080;
PR	30-AUG-1999;	9905-0151303;
PR	31-AUG-1999;	9905-0151438;
PR	01-SEP-1999;	9905-01515930;
PR	07-SEP-1999;	9905-0152363;
PR	10-SEP-1999;	9905-0153070;
PR	13-SEP-1999;	9905-0153578;

PR	15-SEP-1999;	9905-0154018
PR	16-SEP-1999;	9905-0154039
PR	20-SEP-1999;	9905-0154779
PR	22-SEP-1999;	9905-0155133
PR	23-SEP-1999;	9905-0155139
PR	24-SEP-1999;	9905-0155486
PR	28-SEP-1999;	9905-0155659
PR	28-SEP-1999;	9905-0156458
PR	29-SEP-1999;	9905-0156596
PR	04-OCT-1999;	9905-0157117
PR	05-OCT-1999;	9905-0157753
PR	06-OCT-1999;	9905-0157865
PR	07-OCT-1999;	9905-0158029
PR	08-OCT-1999;	9905-0158369
PR	12-OCT-1999;	9905-0158332
PR	13-OCT-1999;	9905-0158293
PR	13-OCT-1999;	9905-0158294
PR	13-OCT-1999;	9905-0158294
PR	13-OCT-1999;	9905-0158295
PR	14-OCT-1999;	9905-0158329
PR	14-OCT-1999;	9905-0158330
PR	14-OCT-1999;	9905-0158331
PR	14-OCT-1999;	9905-0158637
PR	14-OCT-1999;	9905-0158684
PR	18-OCT-1999;	9905-0158588
PR	18-OCT-1999;	9905-0158588
PR	21-OCT-1999;	9905-0160741
PR	21-OCT-1999;	9905-0160767
PR	21-OCT-1999;	9905-0160768
PR	21-OCT-1999;	9905-0160770
PR	21-OCT-1999;	9905-0160814
PR	21-OCT-1999;	9905-0160815
PR	22-OCT-1999;	9905-0160980
PR	22-OCT-1999;	9905-0160981
PR	22-OCT-1999;	9905-0160981
PR	25-OCT-1999;	9905-0161098
PR	25-OCT-1999;	9905-0161404
PR	25-OCT-1999;	9905-0161405
PR	25-OCT-1999;	9905-0161406
PR	26-OCT-1999;	9905-0161359
PR	26-OCT-1999;	9905-0161360
PR	26-OCT-1999;	9905-0161361
PR	28-OCT-1999;	9905-0161920
PR	28-OCT-1999;	9905-0161920
PR	28-OCT-1999;	9905-0161923
PR	29-OCT-1999;	9905-0162142

Query Match	60.3%	Score 1538.5	DB 21	Length 455
Best Local Similarity	61.2%	Pred. No. 5.5e-147		
Matches 279	Conservative 62	Mismatches 112	Indels 3	Gaps 3
QY	19	MKGPREIYLCEIRNNGTEAPDYATVDVDPKSPQCOVTHRLPMPRLKDELHSGWN	78	
Db	1	msgprekilyvaalytgygkqdylatvdvessstysvthrlpmpyldelhlmsgwn	60	
QY	79	TCSSEFGDSTKSRTLYLPLSLISSILYVVDGSEPRAPKLHVIIEPKDINACELALYHT	138	
Db	61	scsccygdscscerrylllpallsgtilyidctlnprepslbfvdpaelectgtlayphq	120	
QY	139	SHCLASGEYMISSLCDVKGNGKG-GEVLIDGETFEVKGTEWBERGGAAPLCYDFWYQPRHN	197	
Db	121	pclslasgdlvscldgedednaegsfllldse-falkgwetkgmnsplgydfwfyqrphk	179	
QY	198	WMSIEMAPANLVRGFPADVEALYXSHLYVMWQRIEYQIOTLSLD-GILPEIRFL	256	
Db	180	tmistswgapaaftkgfdlkdvsdgllygkhlhvywspgqelkqlldgtgtllplevfrfl	239	
QY	257	HNPDAOGFVGALSSSTIORFYKNGKGMVSEKVLQVOPPKKGMGLDEMPMLINDILIS	316	
Db	240	hepdkatgafagaalssctlvrtffkndetweshvaalvepllkwenwllpempjllcdflls	299	
QY	317	LDDELFLYSNMWLGHDRLROYDISDPQRPRLTQGLFGGSIYVKGQPAQVLEDEFLKSQPEPL	376	
Db	300	lddrflfysnmwlghtlrygnledpkrtpvltgslnhvgylvqgkslylalaageekatqfdr	359	
QY	377	VVKGRVAGPOMIOLSDGKRLYTTTSLYSAMDQKQFPDDLREGSVMLQVDDVTVKGL	436	

DB 360 k k i g g r l y g g p m f g l s i d g k r l y v t n s l f s w d r q f g p e l y e k s h m l q i d v d t k g g l 419
QY 437 K L N P N F L V D F G K E P L A P A L A H E L R Y P G D C S S D I W I 472
: |||| |||| | | : |||| : |||| : |||| :
DB 420 s i n p n f v f g t e c p d g p s i a h e m r i p y g g d c t s d i w 455

RESULT 11
AAG36719
ID AAG36719 standard; Protein: 455 AA.
XX AAG36719;
AC
XX 18-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 45039.
DE
XX
KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130447.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131448.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142035.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.

```
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 59.7%: Score 1523.5; DB 21; Length 455;

Best Local Similarity 61.0%: Pred. No. 1.8e-145;

Matches 278; Conservative 64; Mismatches 111; Indels 3; Gaps 3;

```
QY 19 MKGREELVLPICRYRMGFEAPDYATVDPKSPQYCOVTHRLPMPNLKDELHHSGMN 78
Db 1 mspgsekliiytlavtgitgldkpolylatvdgsspsysvinihimpfygdellhsgwn 60
QY 79 TCSSCFDSTKSRKTLVLPSTLISSRIYVVDGSEPRAPKLHKVLEPKDIHAKCELAFLHT 138
```

```
Db 61 scschgdasvdrrylvlpstlisgriyaidtkemprrapslykkyvdpkciadkiglatfpc 120
QY 139 SHCLASGEVMISSLDYKNGKG-GFVILLDETEFEVKGTRPRPGAAPLGDFFYQORHN 197
Db 121 ahclatgelivscldgeegnakngfilldsd-fnkmrwekpbhsplygydfyqdrhk 179
QY 198 VMISTEMAAPNVLRDGFNPADVEAGLVGSHLYVMDQORHETVOTLST-KDGLIFLEIRFL 256
Db 180 tmistwgapkafekfhlqhvadgilyshlnvyswpggeikqjldigpgqllpleirfl 239
QY 257 HNPDAQGFVGCALSTIQREYKNEGTSVEKYIQVPPKKVKGMLLPEMGLITDILLS 316
Db 240 hdpksdgtfygsalsnmirffknsdetshevsvskplkvenwllpempglltfdfls 299
QY 317 LDDREFYFSNMLHGDLDQYDSDPQRLTQQLLGSGIYKGVQVLEDEBLKSQEPFL 376
Db 300 lddrflyfwnlhgdltqynledpknpvltqylwvgslllqkspvkvagdgqntfdfevp 359
QY 377 VVKGRVAGGPOMIQLSIDGKRLYITTSIYSAWDKQFYPPDLIRGYSVMLGVDPVTKGGL 436
Db 360 qikgkslirgpbqmqlsldgkrllyatnslfawdrqfypeimekxshliqdvotekygl 419
QY 437 KLNPNFLVDFGKEPLGPALAHELRYPGDCSSDIWI 472
Db 420 tlnpfdfydgdepdpplsahemrypgdcstsiwi 455

RESULT 12
AAG17316
ID AAG17316 standard; Protein; 455 AA.
XX
AC AAG17316;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18288.
XX
KW Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
```

PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136399.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150864.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152063.
 PR 10-SEP-1999; 99US-0153073.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match	59.48;	Score 1514.5;	DB 21;	Length 455;
Best Local Similarity	60.58;	Pred. No. 1.5e-144;		
Matches 276;	Conservative 65;	Mismatches 112;	Indels 3;	Gaps 3

OY	19	MKPEEELVYLPCJLYRNNGTEAPATLAVNDVDPKSPQCOYIHLPLPMLNKDELLHSGWN	78
Db	1	mspspekilylvatlavylgtgldkprdylatvdvdspspsysvllhlpmpfygdellhsngwn	60
OY	79	TCSGCFGBSTSKRKVLPSLTSSRIYVWDGSEPPARPKLKIVLEPMDIHAKCELAFHT	136
Db	61	sccshcdgasvdrlylvpislsigrlyaidckempredpslykyvdpkeiadktglatphc	120
OY	139	SHCIASGEVMITSLIGDVGKGNGK-GFVILLDGETEFEVKGTWERPGGAAPLGIDEMYOPRH	197
Db	121	ancilatgelilyscljgeegnknkgflllldsd-fniknrwekpahsplygydfwyqprhk	179
OY	198	VMISENMAAPVULRDGFNPADVEAGIVGSHLYVMDWDRHELVORTLS-KOGLJPLELRFL	256
Db	180	tmitstwaprafskfgflghvadqlyshlnhlhyvswwegnekqlldlgpbtldrletfl	233
OY	257	HNPDAOGFEVCALSTIQRFYFKNKGEGTWSEKYIOVPRKKVKMWLPMBGLITYDLIS	316
Db	240	hdpskdctgvtalsalssnmrltffksdetsheavlsyvkrlkvemwllpemgjltdffis	299
OY	317	LDDRRLYSNNLHDDLROYDISDPQRPLTQQLRGSIYVGSRVOYLEDELKSQEPPL	376
Db	300	lddtfiylvnmlhgdlryualiedkcpvltfqyivawgglldqgsbpvkavedgnlfqfevr	355
OY	377	VVKGRVAGGPOMLOLSDGKRKLITYTSLSAMRKOFPYDLIRGSVMLOVDVOTVNGGL	436
Db	360	qikgnslsgygmqlsidsgrkrlyatnslfsawdrqfypelmeagshlqdvotekygl	411
OY	437	KLNPRFLVDFCKEPLGPALAHELKYPGGDCSSDIWT	472
Db	420	tlnpdfitvdigdepelsalemyyggdcscdiwtl	455

RESULT	13
AAG19252	
ID	AAG19252 standard; Protein; 469 AA.

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 20983.

KM Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter
KM termination sequence.

OS *Arabidopsis thaliana*.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR	05-MAR-1999;	99US-0123180.
PR	23 FEB 1999,	99US 0121023.
PR		

PR	09-MAR-1999;	9905-0125348
PR	23-MAR-1999;	9905-0125788
PR	23-MAR-1999;	9905-0126264
PR	29-MAR-1999;	9905-0126785
PR	01-APR-1999;	9905-0127462
PR	06-APR-1999;	9905-0128234
PR	08-APR-1999;	9905-0128714
PR	16-APR-1999;	9905-0129645
PR	19-APR-1999;	9905-0130047
PR	21-APR-1999;	9905-0130479
PR	23-APR-1999;	9905-0130891
PR	28-APR-1999;	9905-0131449
PR	30-APR-1999;	9905-0132048
PR	30-APR-1999;	9905-0132407
PR	04-MAY-1999;	9905-0132484
PR	05-MAY-1999;	9905-0132485
PR	06-MAY-1999;	9905-0132486
PR	07-MAY-1999;	9905-0132487
PR	07-MAY-1999;	9905-0132488
PR	11-MAY-1999;	9905-0132456
PR	11-MAY-1999;	9905-0134218
PR	14-MAY-1999;	9905-0134219
PR	14-MAY-1999;	9905-0134221
PR	14-MAY-1999;	9905-0134370
PR	18-MAY-1999;	9905-0134344
PR	19-MAY-1999;	9905-0134761
PR	20-MAY-1999;	9905-0135124
PR	21-MAY-1999;	9905-0135553
PR	24-MAY-1999;	9905-0135629
PR	25-MAY-1999;	9905-0136021
PR	27-MAY-1999;	9905-0136592
PR	28-MAY-1999;	9905-0136782
PR	01-JUN-1999;	9905-0137222
PR	03-JUN-1999;	9905-0137528
PR	04-JUN-1999;	9905-0137502
PR	07-JUN-1999;	9905-0137724
PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138540
PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139119
PR	16-JUN-1999;	9905-0139452
PR	16-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	18-JUN-1999;	9905-0139462
PR	18-JUN-1999;	9905-0139463
PR	18-JUN-1999;	9905-0139750
PR	18-JUN-1999;	9905-0139760
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0139869
PR	23-JUN-1999;	9905-0140353
PR	24-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140695
PR	28-JUN-1999;	9905-0140823
PR	29-JUN-1999;	9905-0140921
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142390
PR	08-JUL-1999;	9905-0142803
PR	09-JUL-1999;	9905-0142820
PR	12-JUL-1999;	9905-0142977
PR	13-JUL-1999;	9905-0143524
PR	14-JUL-1999;	9905-0143524

XX AC AAG19251;
XX 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 20982.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EPI033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.

```
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 56.6%; Score 1444.5; DB 21; Length 502;
Best Local Similarity 59.7%; Freq. No. 2.2e-137;
Matches 265; Conservative 62; Mismatches 114; Indels 3; Gaps 3;

```
OY 31 CYRRTGTETAPRYLATVNDVDPKSPQYCOYIHRLLPMPNLKDELHNSGMNTSCSGFDSTKS 90
DB 60 cpllngrtdkpylatvdvdpsrpfsvihrlkmpyldelnhgwmnscschngdaaad 119
OY 91 RTKLVPLSISSRIYVVDVSGSEPRAPKLHKYLERPKDIAKCELAFLHNSCLASGEVMS 150
DB 120 rrylvpilpgllgrlyaldtktdkpkarslykvrkerlaektglatrphshncslasgdmlys 179
OY 151 SLGDVKGNGKG-GEVLLDGEFEVKGTEWRPGGAAPLGTDEWYORRHNVMTSEMAAPRV 209
DB 180 clydkegnakngflllisd- fnvksrwdkprhprlfgydfwyqprfkmtstswgapka 238
OY 210 LRDGPNRADVEAGLYGSHLYVMDWQRHELYOFLSL-KDGLRPLELRIERPLHNPRAAGFYGC 268
DB 239 fskgfnlqhavadglyshlyqwrpegemkqlidngntcglrlplelrfllhpskdtgyvgs 298
```

```
OY 269 AISTIQRFYKNEGTSVVEKYIOVPKKYKGLLPEMPGLITDILLSDDRFLYFSMWL 328
DB 299 alsnmirfkrnsdttwshvevisvklkvenmlipempglltdflisldtrfifvfwl 358
OY 329 HCDLRQYDISDPQRRLNGQLFGLSIYKGGVQVLEDEELKSQEPFLVKGKRRVAGGPQ 388
DB 359 hddirgynldpknplvgqlwvsglllqkspykavgednltqfdvppqlykxslragpq 418
OY 389 MQLSLDKRRLYITTSLSAWMDKQFYPPDLIREGSYMLQVDVTVKGLKLNPNFLVDEGK 448
DB 419 mqlsltdgkrllyatnslsawdrqfypeimekxshllqldvtdtkgyltlmpdffvdfgd 478
OY 449 EPLGPALAHLELRYEGGDCSSDIWT 472
DB 479 epdgpalahemrypgdctsdltwl 502

RESULT 15
AAG36720
ID AAG36720 standard; Protein; 409 AA.
XX
AC AAG36720;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 45040.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 09-MAR-1999; 99US-0123180.
PR 23-MAR-1999; 99US-0123548.
PR 25-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0126785.
PR 06-APR-1999; 99US-0127462.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
```

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147392.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155119.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159654.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 29-OCT-1999; 99US-0162142.

Query Match 53.3%; Score 1359.5; DB 21; Length 409;

Best Local Similarity 61.0%; Pred. No. 6,8e-129; Mismatches 99; Indels 3; Gaps 3;

Matches 250; Conservative 58; Mismatches 99; Indels 3; Gaps 3;

```
OY 65 MNLKDELHSGMNTSCSGFDSTKSRKLVLPSSLISRIYVYDVGSSEPRAPKLHKVIEP 124
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 mpfvdelhsgwmscschgdsavdrylviplslsgrlyaidtckenprapslykyvdp 60
OY 125 KDIHAKCELAFLHTSHCLASGEVMISSLDGVNGKKG-GEVLLDGETFEVKGTEWEPGA 183
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 keiacktgtafhtaclaetgelivsclydeegnakngfilllidsd-finknrwekpgfhs 119
OY 184 APLGYDFWYOPRHNWMISTEMAAPNVLRDGFNPADVEAGLYGSHLYWMDQREHIVOTLS 243
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 plgydydfwyqprhkmistswgapkaefskgfnlqhavadglyshlhvyswpggelkqlid 179
OY 244 L-KDGLIPLRFLNHPDAAGFVGCAISSTIQRFYKNEGTVSEKVIQVPPKKVKGWL 302
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 lqptgllpleirflndpskdtgfvgsalsnmirffknsdelcwshevisvxpklkvenwi 239
OY 303 LPEMGLITDILSLDDRFYFSNWLHGDLROYDISDPQRPRLTGOLFSGSIVKGPVQ 362
   |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 lpempglltdflislddrftlyfwnlhgdtrqynledpknpyltgqiwvgllqkyspvk 299
OY 363 VLEDEELKSQPEPLVWKGKRVAGCPQMIQLSIDGKRKYTTTSLYSAMDROFYPPDLIREGS 422
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 avgedgnufqfevpqtkgkslrgpqrmlqslsdgkrlyatnslfawdrqfypeimekys 359
OY 423 VMLQYVDVTVKGGGLKLNPNFLVDFGKEPLGPAHLAHLRYPGDCSSDIWI 472
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 hliqldvdekgylltnpdlfvdgdepdgsiahemrypggdcscsdlwi 409
```

Search completed: August 29, 2002, 16:46:52
Job time: 246 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:44:41 ; Search time 16.33 Seconds
(Without alignments) 705.994 Million cell updates/sec

Title: US-09-841-758-1
Perfect score: 2550
Sequence: 1 MATKCGNCGPGXSTPLEAMK.....PALAHLRYPGDCCSDIWI 472

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCFUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2550	100.0	472	1 US-08-749-903-1	Sequence 1, Appl
2	2550	100.0	472	4 US-09-088-641-1	Sequence 1, Appl
3	2435	95.5	472	1 US-08-749-903-3	Sequence 3, Appl
4	2435	95.5	472	4 US-09-088-641-3	Sequence 3, Appl
5	2280	89.4	472	1 US-08-749-903-5	Sequence 5, Appl
6	2280	89.4	472	4 US-09-088-641-5	Sequence 5, Appl
7	2240	87.8	472	1 US-08-749-903-4	Sequence 4, Appl
8	2240	87.8	472	4 US-09-088-641-4	Sequence 4, Appl
9	102	4.0	543	2 US-08-856-841-22	Sequence 2, Appl
10	98	3.8	365	4 US-09-979-424-3	Sequence 3, Appl
11	98	3.8	365	4 US-09-272-466-2	Sequence 2, Appl
12	95	3.7	672	4 US-08-422-869-2	Sequence 2, Appl
13	95	3.7	757	3 US-08-434-000A-6	Sequence 6, Appl
14	95	3.7	500	4 US-09-312-157-6	Sequence 6, Appl
15	94	3.7	500	4 US-08-856-841-16	Sequence 16, Appl
16	94	3.7	501	4 US-08-856-841-18	Sequence 18, Appl
17	91.5	3.6	264	4 US-08-856-841-14	Sequence 14, Appl
18	91.5	3.6	574	4 US-09-142-623-13	Sequence 13, Appl
19	91	3.6	486	4 US-08-856-841-19	Sequence 19, Appl
20	89.5	3.5	2050	2 US-08-347-594A-2	Sequence 2, Appl
21	89.5	3.5	4472	2 US-08-804-227C-2	Sequence 2, Appl
22	89	3.5	365	4 US-08-928-383B-23	Sequence 23, Appl
23	88	3.5	365	4 US-08-928-383B-26	Sequence 26, Appl
24	87	3.4	415	4 US-08-856-841-20	Sequence 20, Appl
25	87	3.4	911	1 US-08-536-985-2	Sequence 2, Appl
26	86.5	3.4	511	1 US-08-480-604A-20	Sequence 20, Appl
27	86.5	3.4	511	2 US-08-405-496A-20	Sequence 20, Appl

28	86.5	3.4	511	4 US-08-915-136-20	Sequence 20, Appl
29	86.5	3.4	608	1 US-08-480-604A-21	Sequence 21, Appl
30	86.5	3.4	608	2 US-08-405-496A-21	Sequence 21, Appl
31	86.5	3.4	608	4 US-08-915-136-21	Sequence 21, Appl
32	86.5	3.4	609	1 US-08-480-604A-30	Sequence 30, Appl
33	86.5	3.4	609	4 US-08-915-136-30	Sequence 30, Appl
34	86.5	3.4	2366	1 US-08-480-604A-10	Sequence 10, Appl
35	86.5	3.4	2366	2 US-08-405-496A-10	Sequence 10, Appl
36	86.5	3.4	2366	4 US-08-915-136-10	Sequence 10, Appl
37	86	3.4	365	4 US-08-928-383B-2	Sequence 2, Appl
38	86	3.4	365	4 US-08-928-383B-24	Sequence 24, Appl
39	85.5	3.4	1447	4 US-09-041-086-25	Sequence 25, Appl
40	85.5	3.4	1447	5 PCT-US94-05277-2	Sequence 2, Appl
41	85	3.3	779	4 US-09-564-805-235	Sequence 235, App
42	85	3.3	1525	4 US-09-396-651B-1	Sequence 1, Appl
43	84	3.3	612	3 US-09-285-186-16	Sequence 16, Appl
44	83.5	3.3	395	4 US-08-856-841-13	Sequence 13, Appl
45	83.5	3.3	1146	4 US-09-126-109-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-749-903-1
Sequence 1, Application US/08749903
Patent No. 5759812
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: US
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/749,903
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0163 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 472 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5759812e
IMMEDIATE SOURCE:
CLONE: 989953
US-08-749-903-1

Query Match 100.0%; Score 2550; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 1.6e-260;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 MATKCGNGCGPGYSTPLEAMKGPREEIVYLPCYIRNTGTEAPYLATVVDPKSPQYCOVI 60
        |||
        1 MATKCGNGCGPGYSTPLEAMKGPREEIVYLPCYIRNTGTEAPYLATVVDPKSPQYCOVI 60
Db      61 HRLPMPNLKDELHSGMNTCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPKLHK 120
        |||
QY      121 VIEPKDIAKCELAFLHTSHCLASGEVMTISLGDKYGNKGKGFVLLDGETFEVKGTEWRP 180
        |||
Db      121 VIEPKDIAKCELAFLHTSHCLASGEVMTISLGDKYGNKGKGFVLLDGETFEVKGTEWRP 180
        |||
QY      181 GGAAPLGDFWYQPRHNWMTSTEMAAPNVLRDGFNPADVEAGLYGSHLYWMDQREHEIYQ 240
        |||
Db      181 GGAAPLGDFWYQPRHNWMTSTEMAAPNVLRDGFNPADVEAGLYGSHLYWMDQREHEIYQ 240
        |||
QY      241 TSLKDGILPLEIRFLHNDAAQGFVGCALSSSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
        |||
Db      241 TSLKDGILPLEIRFLHNDAAQGFVGCALSSSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
        |||
QY      301 WLLPEMPGLITDILSLDRFLYFSNMHLGDLROYDISDPQPRRLTGOLFSGSYVKGKP 360
        |||
Db      301 WLLPEMPGLITDILSLDRFLYFSNMHLGDLROYDISDPQPRRLTGOLFSGSYVKGKP 360
        |||
QY      361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIQSLDGKRLYITTSLSYSAADKQFYDPLIRE 420
        |||
Db      361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIQSLDGKRLYITTSLSYSAADKQFYDPLIRE 420
        |||
QY      421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLPALAHELRYPGDCSSDIWI 472
        |||
Db      421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLPALAHELRYPGDCSSDIWI 472
        |||

```

RESULT 2

```

US-09-088-641-1
; Sequence 1, Application us/09088641
; Patent No. 6312895
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,641
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/749,903
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: No. 6312895e
; IMMEDIATE SOURCE:
; CLONE: 989953
US-09-088-641-1

```

```

Query Match      100.0%; Score 2550; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 1,6e-260;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MATKCGNGCGPGYSTPLEAMKGPREEIVYLPCYIRNTGTEAPYLATVVDPKSPQYCOVI 60
        |||
Db      1 MATKCGNGCGPGYSTPLEAMKGPREEIVYLPCYIRNTGTEAPYLATVVDPKSPQYCOVI 60
        |||
QY      61 HRLPMPNLKDELHSGMNTCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPKLHK 120
        |||
Db      61 HRLPMPNLKDELHSGMNTCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPKLHK 120
        |||
QY      121 VIEPKDIAKCELAFLHTSHCLASGEVMTISLGDKYGNKGKGFVLLDGETFEVKGTEWRP 180
        |||
Db      121 VIEPKDIAKCELAFLHTSHCLASGEVMTISLGDKYGNKGKGFVLLDGETFEVKGTEWRP 180
        |||
QY      181 GGAAPLGDFWYQPRHNWMTSTEMAAPNVLRDGFNPADVEAGLYGSHLYWMDQREHEIYQ 240
        |||
Db      181 GGAAPLGDFWYQPRHNWMTSTEMAAPNVLRDGFNPADVEAGLYGSHLYWMDQREHEIYQ 240
        |||
QY      241 TSLKDGILPLEIRFLHNDAAQGFVGCALSSSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
        |||
Db      241 TSLKDGILPLEIRFLHNDAAQGFVGCALSSSTIQRFYKNEGTSVEKVIQVPPKRYKG 300
        |||
QY      301 WLLPEMPGLITDILSLDRFLYFSNMHLGDLROYDISDPQPRRLTGOLFSGSYVKGKP 360
        |||
Db      301 WLLPEMPGLITDILSLDRFLYFSNMHLGDLROYDISDPQPRRLTGOLFSGSYVKGKP 360
        |||
QY      361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIQSLDGKRLYITTSLSYSAADKQFYDPLIRE 420
        |||
Db      361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIQSLDGKRLYITTSLSYSAADKQFYDPLIRE 420
        |||
QY      421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLPALAHELRYPGDCSSDIWI 472
        |||
Db      421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLPALAHELRYPGDCSSDIWI 472
        |||

```

RESULT 3

```

US-08-749-903-3
; Sequence 3, Application US/08749903
; Patent No. 5759812
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,903
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

```



```

; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 5759812e
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 374792
US-08-749-903-3

Query Match          95.5%; Score 2435; DB 1; Length 472;
Best Local Similarity 96.0%; Pred. No. 2.2e-248;
Matches 453; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

QY 1 MATCGNGPGYSTPLAMKGPREEIYLPCTYRNTGTEADYATADVDPKSPQYCYI 60
Db 1 MATCGNGPGYSTPLAMKGPREEIYLPCTYRNTGTEADYATADVDPKSPQYCYI 60
QY 61 HRLPMNLKDELHSGNNTSCSCFGDSTKSRKLVLPSSLISRIYVDVSGSEPPQKLHK 120
Db 61 HRLPMNLKDELHSGNNTSCSCFGDSTKSRKLVLPSSLISRIYVDVSGSEPPQKLHK 120
QY 121 VIEPKDIHAKCELAFLHTSHCLASGEVWISSLGDKNGKGFVLLDGETFEVKGWERP 180
Db 121 VIEPKDIHAKCELAFLHTSHCLASGEVWISSLGDKNGKGFVLLDGETFEVKGWERP 180
QY 181 GGAAPLGVDFEYQORHNVMISTEWAAPNVLRDGNPADVEAGLGSHTLYWDMQORHEIYQ 240
Db 181 GGAAPLGVDFEYQORHNVMISTEWAAPNVLRDGNPADVEAGLGSHTLYWDMQORHEIYQ 240
QY 241 TLSLKDGILPLEIRFLNPDAAQGFVGCALSTIQREYKKNEGTWSVEKYIQVPPKKYVG 300
Db 241 TLSLKDGILPLEIRFLNPDAAQGFVGCALSTIQREYKKNEGTWSVEKYIQVPPKKYVG 300
QY 301 WLPFMGGLITDILSLDDREFLYFSNNLHGDRLROYDISDPQRPRLTQQLFLGGSIVKGGP 360
Db 301 WLPFMGGLITDILSLDDREFLYFSNNLHGDRLROYDISDPQRPRLTQQLFLGGSIVKGGP 360
QY 361 VOVEDEELKSQPEPLVYKGRVAGGPMQIOLSLDGKRLYTTSTLYSAMDKQFPDLIRE 420
Db 361 VOVEDEELKSQPEPLVYKGRVAGGPMQIOLSLDGKRLYTTSTLYSAMDKQFPDLIRE 420
QY 421 GSVMLQYDVDTYVKGGLKLNPNFLVDFGKEPLGPAALAHELRYPGDCCSSDIWI 472
Db 421 GSVMLQYDVDTYVKGGLKLNPNFLVDFGKEPLGPAALAHELRYPGDCCSSDIWI 472

RESULT 4
US-09-088-641-3
; Sequence 3, Application us/09088641
; Patent No. 6312895
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/088,641
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/749,903
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6312895e
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 374792
US-09-088-641-3

Query Match          95.5%; Score 2435; DB 4; Length 472;
Best Local Similarity 96.0%; Pred. No. 2.2e-248;
Matches 453; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

QY 1 MATCGNGPGYSTPLAMKGPREEIYLPCTYRNTGTEADYATADVDPKSPQYCYI 60
Db 1 MATCGNGPGYSTPLAMKGPREEIYLPCTYRNTGTEADYATADVDPKSPQYCYI 60
QY 61 HRLPMNLKDELHSGNNTSCSCFGDSTKSRKLVLPSSLISRIYVDVSGSEPPQKLHK 120
Db 61 HRLPMNLKDELHSGNNTSCSCFGDSTKSRKLVLPSSLISRIYVDVSGSEPPQKLHK 120
QY 121 VIEPKDIHAKCELAFLHTSHCLASGEVWISSLGDKNGKGFVLLDGETFEVKGWERP 180
Db 121 VIEPKDIHAKCELAFLHTSHCLASGEVWISSLGDKNGKGFVLLDGETFEVKGWERP 180
QY 181 GGAAPLGVDFEYQORHNVMISTEWAAPNVLRDGNPADVEAGLGSHTLYWDMQORHEIYQ 240
Db 181 GGAAPLGVDFEYQORHNVMISTEWAAPNVLRDGNPADVEAGLGSHTLYWDMQORHEIYQ 240
QY 241 TLSLKDGILPLEIRFLNPDAAQGFVGCALSTIQREYKKNEGTWSVEKYIQVPPKKYVG 300
Db 241 TLSLKDGILPLEIRFLNPDAAQGFVGCALSTIQREYKKNEGTWSVEKYIQVPPKKYVG 300
QY 301 WLPFMGGLITDILSLDDREFLYFSNNLHGDRLROYDISDPQRPRLTQQLFLGGSIVKGGP 360
Db 301 WLPFMGGLITDILSLDDREFLYFSNNLHGDRLROYDISDPQRPRLTQQLFLGGSIVKGGP 360
QY 361 VOVEDEELKSQPEPLVYKGRVAGGPMQIOLSLDGKRLYTTSTLYSAMDKQFPDLIRE 420
Db 361 VOVEDEELKSQPEPLVYKGRVAGGPMQIOLSLDGKRLYTTSTLYSAMDKQFPDLIRE 420
QY 421 GSVMLQYDVDTYVKGGLKLNPNFLVDFGKEPLGPAALAHELRYPGDCCSSDIWI 472
Db 421 GSVMLQYDVDTYVKGGLKLNPNFLVDFGKEPLGPAALAHELRYPGDCCSSDIWI 472

RESULT 5
US-08-749-903-5
; Sequence 5, Application us/08749903
; Patent No. 5759812
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
```

```

? NUMBER OF SEQUENCES: 5
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
? STREET: 3174 Porter Drive
? CITY: Palo Alto
? STATE: CA
? COUNTRY: US
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/749,903
? FILING DATE: Filed Herewith
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Billings, Lucy J.
? REGISTRATION NUMBER: 36,749
? REFERENCE/DOCKET NUMBER: PF-0163 US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-855-0555
? TELEFAX: 415-845-4166
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 472 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? IMMEDIATE SOURCE:
? LIBRARY: Genbank
? CLONE: 298710
?
? US-08-749-903-5

```

```

Query Match      89.4%; Score 2280; DB 1; Length 472;
Best Local Similarity 87.7%; Pred. No. 5,1e-232;
Matches 414; Conservative 27; Mismatches 31; Indels 0; Gaps 0;

```

```

QY 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVVDPKSPQYCOVI 60
DB 1 MATKCTKCGPGSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVVDPKSPQYCOVI 60
QY 61 HRLPMPNLKDELHSHGWNMTSCSFCGDSRTKSRKIVLPGLISSRIYVVDGSEPRAPKXHK 120
DB 61 HRLPMPNLKDELHSHGWNMTSCSFCGDSRTKSRKIVLPGLISSRIYVVDGSEPRAPKXHK 120
QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLGDKVKGNGKGGEVLLDGETFEVKGTEWERP 180
DB 121 VIEASITQAKCNVSNHTSHCLASGEVMTSSLGDKVKGNGKGGEVLLDGETFEVKGTEWERP 180
QY 181 GGAAPLGIDFWYQPRHNWMTSTEWAAPNVLRDGFNPADYVAGLYSHLYVMDMORHEITQ 240
DB 181 GASPMPGYDFWYQPRHNWMTSTEWAAPNVLRDGFNPADYVAGLYSHLYVMDMORHEITQ 240
QY 241 TSLKDGILPLEIRFLHNPDAAGFVGALSSITQRFYKNEGGTWSEVKVIOVPPSKKYKG 300
DB 241 TLOMTDGLPLERFLHNPDAAGFVGALSSITQRFYKNEGGTWSEVKVIOVPPSKKYKG 300
QY 301 WLLPEMGLITDILSLDRFLYFSNWLHGLDROYDISDPORPRLTGOFLGGSIVYKGP 360
DB 301 WMLPEMGLITDILSLDRFLYFSNWLHGLDROYDISDPORPRLTGOFLGGSIVYKGP 360
QY 361 VOYLEDEELKSPREPLVYVKGKRVAGPQMTQLSLDGKRLYITTSYLSAMDKQFYDPLIRE 420
DB 361 VOYLEDEELKSPREPLVYVKGKRVAGPQMTQLSLDGKRLYITTSYLSAMDKQFYDPLIRE 420
QY 421 GSVMLQVNDVDTYKGLKLPNPLVNDGKEPLPALAHELRVYGGCCSSDIWI 472
DB 421 GSVMLQVNDVDTYKGLKLPNPLVNDGKEPLPALAHELRVYGGCCSSDIWI 472

```

```

DB 421 GSVMLQVNDVDTYKGLKLPNPLVNDGKEPLPALAHELRVYGGCCSSDIWI 472

```

```

RESULT 6
US-09-088-641-5
? Sequence 5, Application US/09088641
? Patent No. 6312895
? GENERAL INFORMATION:
? APPLICANT: Bandman, Olga
? APPLICANT: Hawkins, Phillip R.
? TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
? NUMBER OF SEQUENCES: 5
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
? STREET: 3174 Porter Drive
? CITY: Palo Alto
? STATE: CA
? COUNTRY: US
? ZIP: 94304
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette
? COMPUTER: IBM Compatible
? OPERATING SYSTEM: DOS
? SOFTWARE: FASTSEQ Version 1.5
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/088,641
? FILING DATE:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/749,903
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Billings, Lucy J.
? REGISTRATION NUMBER: 36,749
? REFERENCE/DOCKET NUMBER: PF-0163 US
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-855-0555
? TELEFAX: 415-845-4166
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 472 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? IMMEDIATE SOURCE:
? LIBRARY: Genbank
? CLONE: 298710
?
? US-09-088-641-5

```

```

Query Match      89.4%; Score 2280; DB 4; Length 472;
Best Local Similarity 87.7%; Pred. No. 5,1e-232;
Matches 414; Conservative 27; Mismatches 31; Indels 0; Gaps 0;

```

```

QY 1 MATKCGNGCGPGSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVVDPKSPQYCOVI 60
DB 1 MATKCTKCGPGSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVVDPKSPQYCOVI 60
QY 61 HRLPMPNLKDELHSHGWNMTSCSFCGDSRTKSRKIVLPGLISSRIYVVDGSEPRAPKXHK 120
DB 61 HRLPMPNLKDELHSHGWNMTSCSFCGDSRTKSRKIVLPGLISSRIYVVDGSEPRAPKXHK 120
QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSSLGDKVKGNGKGGEVLLDGETFEVKGTEWERP 180
DB 121 VIEASEITQAKCNVSNHTSHCLASGEVMTSSLGDKVKGNGKGGEVLLDGETFEVKGTEWERP 180
QY 181 GGAAPLGIDFWYQPRHNWMTSTEWAAPNVLRDGFNPADYVAGLYSHLYVMDMORHEITQ 240
DB 181 GASPMPGYDFWYQPRHNWMTSTEWAAPNVLRDGFNPADYVAGLYSHLYVMDMORHEITQ 240
QY 241 TSLKDGILPLEIRFLHNPDAAGFVGALSSITQRFYKNEGGTWSEVKVIOVPPSKKYKG 300
DB 241 TLOMTDGLPLERFLHNPDAAGFVGALSSITQRFYKNEGGTWSEVKVIOVPPSKKYKG 300

```

QY 301 WLPEMPLITDILSLDDRELYFSNMHGDRLROYDISDQRPRLTGOLFSGSIYVGGP 360
 Db 301 WLPEMPLITDILSLDDRELYFSNMHGDRLROYDISDQRPRLTGOLFSGSIYVGGP 360
 QY 361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIOQLSDGKRLYTTSLSYAMDQKOFYDILRE 420
 Db 361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIOQLSDGKRLYTTSLSYAMDQKOFYDILRE 420
 QY 421 GSVMLQYDVDTYVKGKRLNPNFLVDGKPEPLGAPLAHELRYPGGDCSSDIWI 472
 Db 421 GSVMLQYDVDTYVKGKRLNPNFLVDGKPEPLGAPLAHELRYPGGDCSSDIWI 472

RESULT 7
 US-08-749-903-4
 ; Sequence 4, Application US/08749903
 ; Patent No. 5759812
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/749,903
 ; FILING DATE: Filed Herewith
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0163 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 472 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 227630
 ; US-08-749-903-4

Query Match 87.8%; Score 2240; DB 1; Length 472;
 Best Local Similarity 86.0%; Pred. No. 8.6e-228;
 Matches 406; Conservative 29; Mismatches 37; Indels 0; Gaps 0;

QY 1 MATKCGMGCGYSTPLEAMGPREELIYLCITYRNTGTEADPIYATVDPKSPQYCVI 60
 Db 1 MATKCTCGGCGYSTPLEAMGPREELIYLCITYRNTGTEADPIYATVDPKSPQYCVI 60
 QY 61 HRLPMPLKDELHSHGNTSCSGDSTKSRTKLVLPSLISRYIVYVDSPPAPRLHK 120
 Db 61 HRLPMPLKDELHSHGNTSCSGDSTKSRTKLVLPGLISSRIYIVYVDSPPAPRLHK 120

QY 121 VIEPKDIHAKCELAFLTSHCLASGEYVWISLGDVKGNGKGGFVLLDGETFEVKGTEWRP 180
 Db 121 VIEASEIQAKCNVSSLHTSHCLASGEYVWSTLGIQNGKGSFVLLDGETFEVKGTEWRP 180
 QY 181 GGAAPLGYDEWYQPRHNVMISTEWAAPNVLNDGFNPADVEAGLYSHLYWDMQRHEIVQ 240
 Db 181 GDAAPMGYDEWYQPRHNVMISTEWAAPNVFNKDNPAHVEAGLYSGRLFWDMQRHEIIQ 240
 QY 241 TLSLKGILPLEIFLNPDAAGFVCCALSTIQRYKKNKGGSVEKYIYVPPKRYVG 300
 Db 241 TLMNTDGLPLEIFLNDPSATQGVGCASAPNIOREYKNAEGTWSVEKYIYVPPKRYVG 300
 QY 301 WLPEMPLITDILSLDDRELYFSNMHGDRLROYDISDQRPRLTGOLFSGSIYVGGP 360
 Db 301 WLPEMPLITDILSLDDRELYFSNMHGDRLROYDISDQRPRLTGOLFSGSIYVGGP 360
 QY 361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIOQLSDGKRLYTTSLSYAMDQKOFYDILRE 420
 Db 361 VOYLEDEELKSQPEPLVYVKGKRVAGPQMIOQLSDGKRLYTTSLSYAMDQKOFYDILRE 420
 QY 421 GSVMLQYDVDTYVKGKRLNPNFLVDGKPEPLGAPLAHELRYPGGDCSSDIWI 472
 Db 421 GSVMLQYDVDTYVKGKRLNPNFLVDGKPEPLGAPLAHELRYPGGDCSSDIWI 472

RESULT 8
 US-09-088-641-4
 ; Sequence 4, Application US/09088641
 ; Patent No. 6312895
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FASTSEQ Version 1.5
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/088,641
 ; FILING DATE:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/749,903
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0163 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-855-0555
 ; TELEFAX: 415-845-4166
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 472 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GenBank
 ; CLONE: 227630
 ; US-09-088-641-4

Query Match 87.8%; Score 2240; DB 4; Length 472;

Best Local Similarity 86.0%; Pred. No. 8.6e-228;
Matches 406; Conservative 29; Mismatches 37; Indels 0; Gaps 0;

QY	1	MATKONGSPYSTRPLEMKGPREIYILPCITYKNTSTEARPDYATADVPKSPQYCVI	60
Db	1	MATKCTKGPSTPPLKMKPRREIYILPCITYKNTSTEARPDYATADVPKSPQYSQVI	60
QY	61	HLPLPMLKDELHHSWNTCSSCFGSDSTKSLVLPDLSSRIYVVDGSEPRAPKLHK	120
Db	61	HLPLPPLYLKDELHHSWNTCSSCFGSDSTKSNKLLILGELISSRIYVVDGSEPRAPLHK	120
QY	121	VIEPKDIHAKCELFLYHSHCLAGEVWISLGVKNKGSGFLLDGEFFEVKGWTERP	180
Db	121	VIEASEIDAKCNVSSLATHSLCLAGEVWVSTLGIQSGKSGFLLDGETFEVKGWTEKP	180
QY	181	GGAAPLGYDFWYQOPRHNNMISTEAAAPNVLDGFPNPAVDENAGLGSYHLYWDMQRIETVQ	240
Db	181	GDAAPMGVDFWYQOPRHNNMISTEAAAPNVFEDGFPNPAHVEAGLGSYHLPWMDQRHILIQ	240
QY	241	TLSLKDGLIPLEIRLPANPDAAQGVGCAALSTIQRFTKKNBGTSVYEVKYIYVPPKKYVG	300
Db	241	TLOMTDGLIPLEIRFLHPDPSATQGVGASAPNQRFTKKNBGTSVYEVKYIYVPPSKYVG	300
QY	301	WLLPEMPELITDILSLDDRFLYKSNMLHGOBLRQYDLSDDPRPLTGOLFJGSGIVYKGP	360
Db	301	WMLPEVPELITDILSLDDRFLYKSNMLHGDITRQYDLSNPQKPLPAQOIFLGSGIVYKGS	360
QY	361	VOVLEDEBLKSQPEPLVYKGRVAGGPOMIOLSLDGKRLYITTSLYSAMDKQFYPDILRE	420
Db	361	VOVLEDEGLTQPEPLVYKGRIRPGGPOMIOLSLDGKRLYATTSLYSAMDKQFYPDILRE	420
QY	421	GSWMQYVVDYVYKGLKLNPNFVLDGFEKPELGPALHAHLKRPGGDCSSDIWI	472
Db	421	GSMMQYDIDYVYVNGGLKLNPNFVLDGFLPGALHAHLKRPGGDCSSDIWI	472

RESULT 9
 US-08-856-841-22
 ; Sequence 22, Application US/08856841
 ; Patent No. 6274307
 ; GENERAL INFORMATION:
 APPLICANT: ERWIN SOUTSCHEK
 APPLICANT: MANFRED MÖTZ
 TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
 TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: BROOKS HAIDT HAEFFNER & DELAHUNTY
 STREET: 99 PARK AVENUE
 CITY: NEW YORK
 STATE: NY
 COUNTRY: USA
 ZIP: 10016
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" FLOPPY DISC
 COMPUTER: AT&T - IBM COMPATIBLE
 OPERATING SYSTEM: MS-DOS Version 6.2
 SOFTWARE: ASCII
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/856,841
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/214,658
 FILING DATE: 16-MARCH-1994
 APPLICATION NUMBER: US 07/917,096
 FILING DATE: 4-AUGUST-1992
 APPLICATION NUMBER: PCT/DE91/00106
 FILING DATE: 8-FEBRUARY-1991
 APPLICATION NUMBER: DE40036262
 FILING DATE: 8-FEBRUARY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: ROBINSON, WILLIAM R.

REGISTRATION NUMBER: 27, 224
REFERENCE/DOCKET NUMBER: LKR-9922-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 697-3355
TELEFAX: (212) 557-5635
TELEX: NONE
INFORMATION FOR SEQ. ID NO: 22:
SEQUENCE CHARACTERISTICS:

```

Query Match          4.0% ; Score 102 ; DB 4 ; Length 543 ;
Best Local Similarity 22.9% ; Pred. No. 0.056 ;
Matches 65 ; Conservative 29 ; Mismatches 80 ; Indels 110 ; Gaps 13 ;

QY      11 GYSTPLEAMKKGREIIVLYLPC-----IYRNTGAEADYL-----ATVDVDRKSPQYCOV 59
      ||||| : : : : : ||| ||| : ||| : ||
Db      91 GISTPWRVYIDFNALNIFSPLEFOHLIENYGSIAIDALTYTISELAVADYDTKRTGGYVY 150
      ||||| : : : : : ||| ||| : ||| : ||

QY      60 IIRLPMNLKDELHHSWNVTCSCSGDSTKSRKTLVPLSLISSRI---YVVDVGSEPRAP 116
      ||| : : : : : ||| : : : : :
Db      151 T-----DSTTGR-----LCMLVDHEKRYRVLVGGQDTLAP 181
      ||| : : : : : ||| : : : : :

QY      117 KLHVIPEPKDIHAKCELAFLLHTSHCLASGEVNISSIGDYKNGKGG-----FVL 165
      || : : : : : ||| : : : : :
Db      182 EL-----PIWVFFPPQYAL-----IYGDVNTGIGISGDSKKLASESAFY 222
      || : : : : : ||| : : : : :

QY      166 LDGETFEVYKGIWERBGGAAFLGYPDWYORRHNVMISTEMAANVLRDGFNPADVAGLYG 225
      || : : : : : ||| : : : : :
Db      223 LEHSSFQLLGT---GGTASMSYKPPRPVRENLE-----GCSQHLYEMNP-----LYG 267
      || : : : : : ||| : : : : :

QY      226 SHLYVWDV-----QRHE-----IVOTLSLKDQ 247
      ||| : : : : : ||| : : : : :
Db      268 SRLGVDPDLTGDPKFRSLTHEDHATQPNQMPGRLVNASYSTKEG 311
      ||| : : : : : ||| : : : : :

RESULT 10
US-08-979-424-3
; Sequence 3, Application US/08979424
; Patent No. 5942606

```

GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: VIRAL RECEPTOR PROTEIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/979,424
FILING DATE: Filed Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0405 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Genbank
CLONE: 1946351
US-08-979-424-3

Query Match 3.8%; Score 98; DB 2; Length 365;
Best Local Similarity 17.4%; Pred. No. 0.081;
Matches 82; Conservative 72; Mismatches 140; Indels 178; Gaps 21;

13 STPLEAMKGPREEIVYLPCIRNMTGEAPDYLATVVD-PKSPQCYVIRLPM----- 65
Db 23 TTPEMIEKAKGETAYLPCKF-----TLPEDQGLDIEMLISPADNQKVDQVILYSGDK 78
QY 66 -----PILKDELHSGWNTSCSGFDSTKRTLVLPSSLISRIYVDVGS-----EPRA 115
Db 79 IYDDYDPLKGRVHFTS-NDLKS--GDASINVTNQLS-----DIGTYCKYKKA 125
QY 116 P-----KLHKVIEPKDIHAKCELAFLHTSHCLASGEVWISSLDGVKNGKGGFVLLDGET 170
Db 126 PGVANKKIHLYLVKPPSGARC-----YDGSSEI-----GSD 157
QY 171 FEVKGTEWRRPGGAAPLGDFWYQPRHNVMISTEAAAPVLRDGFNPADVEAGLYGSHLYV 230
Db 158 FKIR--CEPKRGSLPLOYE-WQKLSDSOKMPTSWLA----- 190
QY 231 WDMORHEIVOT-LSLKDGLPLEIRFLHNPDAAGFVGCALSTIQRFYKNEGTSWYEVK 289
Db 191 -----EXTSSVISYKNA-----SSEYGTYSCTVRNRY-----GSDOCLLR 226
QY 290 VIOVPPKKVKGWMLPEMGLITDILSL-----DREFLYSNMLHGLRDYDI 337
Db 227 LNVVPPSKKAGLI---AGALITGLLALALIGLIIPCCKRRKREKEVEVHHIRE-DV 281
QY 338 SDPORPRLTGQLFLGSGIVKGPV-----QVLEDELKSQPEPLVVKGRVA 384

13 STPLEAMKGPREEIVYLPCIRNMTGEAPDYLATVVD-PKSPQCYVIRLPM----- 65
Db 23 TTPEMIEKAKGETAYLPCKF-----TLPEDQGLDIEMLISPADNQKVDQVILYSGDK 78
QY 66 -----PILKDELHSGWNTSCSGFDSTKRTLVLPSSLISRIYVDVGS-----EPRA 115
Db 79 IYDDYDPLKGRVHFTS-NDLKS--GDASINVTNQLS-----DIGTYCKYKKA 125
QY 116 P-----KLHKVIEPKDIHAKCELAFLHTSHCLASGEVWISSLDGVKNGKGGFVLLDGET 170
Db 126 PGVANKKIHLYLVKPPSGARC-----YDGSSEI-----GSD 157
QY 171 FEVKGTEWRRPGGAAPLGDFWYQPRHNVMISTEAAAPVLRDGFNPADVEAGLYGSHLYV 230
Db 158 FKIR--CEPKRGSLPLOYE-WQKLSDSOKMPTSWLA----- 190
QY 231 WDMORHEIVOT-LSLKDGLPLEIRFLHNPDAAGFVGCALSTIQRFYKNEGTSWYEVK 289
Db 191 -----EXTSSVISYKNA-----SSEYGTYSCTVRNRY-----GSDOCLLR 226
QY 290 VIOVPPKKVKGWMLPEMGLITDILSL-----DREFLYSNMLHGLRDYDI 337
Db 227 LNVVPPSKKAGLI---AGALITGLLALALIGLIIPCCKRRKREKEVEVHHIRE-DV 281
QY 338 SDPORPRLTGQLFLGSGIVKGPV-----QVLEDELKSQPEPLVVKGRVA 384
Db 342 A-----PILSRMGALIPVMIIPAQSKDGI 364

RESULT 11
US-09-272-496-2
Sequence 2, Application US/09272496
Patent No. 6245966
GENERAL INFORMATION:
APPLICANT: Degregori, James
TITLE OF INVENTION: Adenoviral mediated gene transfer into lymphocytes
FILE REFERENCE: 90-98
CURRENT APPLICATION NUMBER: US/09/272,496
CURRENT FILING DATE: 1999-03-19
EARLIER APPLICATION NUMBER: US 60/092782
EARLIER FILING DATE: 1998-07-14
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 2
LENGTH: 365
TYPE: PRT
ORGANISM: Homo sapiens
US-09-272-496-2

Query Match 3.8%; Score 98; DB 4; Length 365;
Best Local Similarity 17.4%; Pred. No. 0.081;
Matches 82; Conservative 72; Mismatches 140; Indels 178; Gaps 21;

13 STPLEAMKGPREEIVYLPCIRNMTGEAPDYLATVVD-PKSPQCYVIRLPM----- 65
Db 23 TTPEMIEKAKGETAYLPCKF-----TLPEDQGLDIEMLISPADNQKVDQVILYSGDK 78
QY 66 -----PILKDELHSGWNTSCSGFDSTKRTLVLPSSLISRIYVDVGS-----EPRA 115
Db 79 IYDDYDPLKGRVHFTS-NDLKS--GDASINVTNQLS-----DIGTYCKYKKA 125
QY 116 P-----KLHKVIEPKDIHAKCELAFLHTSHCLASGEVWISSLDGVKNGKGGFVLLDGET 170
Db 126 PGVANKKIHLYLVKPPSGARC-----YDGSSEI-----GSD 157
QY 171 FEVKGTEWRRPGGAAPLGDFWYQPRHNVMISTEAAAPVLRDGFNPADVEAGLYGSHLYV 230
Db 158 FKIR--CEPKRGSLPLOYE-WQKLSDSOKMPTSWLA----- 190
QY 231 WDMORHEIVOT-LSLKDGLPLEIRFLHNPDAAGFVGCALSTIQRFYKNEGTSWYEVK 289
Db 191 -----EXTSSVISYKNA-----SSEYGTYSCTVRNRY-----GSDOCLLR 226
QY 290 VIOVPPKKVKGWMLPEMGLITDILSL-----DREFLYSNMLHGLRDYDI 337
Db 227 LNVVPPSKKAGLI---AGALITGLLALALIGLIIPCCKRRKREKEVEVHHIRE-DV 281
QY 338 SDPORPRLTGQLFLGSGIVKGPV-----QVLEDELKSQPEPLVVKGRVA 384
Db 385 GGPQMQLSDGKRRLYITTSLSAMDQOFYDPLIREGSVMLQVVDVVGGL 436
Db 342 A-----PILSRMGALIPVMIIPAQSKDGI 364

RESULT 12
US-09-422-869-2
Sequence 2, Application US/09422869
Patent No. 6235481
GENERAL INFORMATION:
APPLICANT: POLONSKY, KENNETH S.
APPLICANT: HORIKAWA, YUKIO

```

; APPLICANT: ODA, NAOHISA
; APPLICANT: COX, NANCY J.
; APPLICANT: SREENAN, SEAMUS
; APPLICANT: ZHOU, YUN-PING
; APPLICANT: OTANI, KENICHI
; APPLICANT: HANIS, CRAIG L.
; APPLICANT: BELL, GRAEME I.
; TITLE OF INVENTION: METHODS OF TREATMENT OF TYPE 2 DIABETES
; FILE REFERENCE: ARCD:307
; CURRENT APPLICATION NUMBER: US/09/422,869
; CURRENT FILING DATE: 1999-10-21
; EARLIER APPLICATION NUMBER: 60/134,175
; EARLIER FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 672
; TYPE: PRT
; ORGANISM: Human
; US-09-422-869-2

```

```

Query Match          3.7%; Score 95; DB 4; Length 672;
Best Local Similarity 20.8%; Pred. No. 0.45;
Matches 69; Conservative 44; Mismatches 134; Indels 84; Gaps 14;

```

```

OY 49 VDPKSPQYCOVTHRLPMLKDELHHSQW-NFCCSCFGDSTKRTKLYLP---SLISSR 103
DB 371 VSEPESEVYIAVQIR-----SRHAAADNAGARALVGSHTSWSPASITPGHXYQAVGH 423
OY 104 IYVDVGSSEPPAKLHKVIE-----PKDIHAKCELA-----FLHTS-- 139
DB 424 LMKV-----EKRRVNIPLRYLSMPVPVAGTACHAYDREVHLRCELSPGYLLAVPSTFLKADPG 479
OY 140 ----HCLASGEWISLGVKNGKGGFVLDE--TEFVKGTWERPAGGAALGDFEWQ 193
DB 480 EELLRYFSTGRVSLSKIRAVAKNTTPGAALPAGEWGTQLRGSW-RVGTQAGGSRNFIASY 538
OY 194 PRHNWIMISTEMAAPN-----VLRDGFNPADVEAGLYGSHLY-VMDWRHEIVQTLSLKD 246
DB 539 PINPCPEFVSPGPGRCVRYRITLHQCRRSPDTEPHDIGHFIOVPEGGRSQDAPRLLOE 598
OY 247 GLIPLERLHPDAAGVGCALSTIORFYKNEGTSSVEKVIQVPP---KRYKGMLL 303
DB 599 PLL-----SCVPHRYAOE-----VSRLLPLPAGTYKVPSTYL 631
OY 304 PEMPGILITDILSLDRFLYFSNMLHGLDRO 334
DB 632 PDEGAFYVTAIRIDRPSIHQEMLGRLO 662

```

```

RESULT 13
US-08-434-000A-6
; Sequence 6, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 Mb
; COMPUTER: IBM Compatible

```

```

; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: Including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 757 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-434-000A-6
; Bovine Polymunoglobulin Receptor

```

```

Query Match          3.7%; Score 95; DB 3; Length 757;
Best Local Similarity 22.0%; Pred. No. 0.55;
Matches 88; Conservative 51; Mismatches 125; Indels 136; Gaps 26;

```

```

OY 4 KCG-----NCGPESTYPLEAMKPRE-----ETVYLPCTIYRMTGTPADYLATV 47
DB 109 KCGLTISSRGLNFDVSLVSDPPADASHAHVYITDILGRVITNCFPTANSKRKSLC-- 166
OY 48 DVDKSPQYC-OVHIR-LPMLKDELHHS--GNNTCCSCFGDSTKRTKLYLPSSLSS 102
DB 167 ---KRTIDCCFQVVDSTGVSNYSKRAHISILGTNT-----LVESVAIN 208
OY 103 RI-----YVVDGSEPPAKLH---KYIEPK-----DIHAK-----CELA----- 134
DB 209 RKLSDAGNYVQOADDKADKINIDLOYLEPELVLVYGDLSVTFPDSLGEPVANYPK 268
OY 135 FLHTSHCLASGEWISLGVKNGKGGFVL--DGETFEVKGWERPQ-----GAAP 185
DB 269 FLQCKKNGGACAVYINTLGRKADPOGRIVSPKONGVFSVNIITSLRKEDAGRYVCGAQ 328
OY 186 LG-YDFWYOPRHNWIMISTEMAAP-----NVLRDG-----FNPADVEAGLYGSHLYV 230
DB 329 EGEPPDGMFVQAMOLFVNEETAIIPASPSVYKGVSGSVTSCPNPKDANSKRYMCH-- 385
OY 231 WDMQNH-----ETVQTLSLK-----GGLIPLERLHPDAAGVGCALG 271
DB 386 WEEAONGRCPLVESRGLMKKEQYEGRLVLLTEPGNGYTVIINQLDDDA--GTYWCV-- 441
OY 272 STIQFYKNKGGT-W--SVE-KVIO-----VPPKRVKGM 302
DB 442 -----TDGDTRWISTVELKVKVGGSPSLKVPKNVTAML 473

```

```

RESULT 14
US-09-312-157-6
; Sequence 6, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19

```

;;
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Lyon & Lyon
;; STREET: 633 West Fifth Street
;; Suite 4700
;; CITY: Los Angeles
;; STATE: California
;; COUNTRY: U.S.A.
;; ZIP: 90071
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
;;
;; STORAGE
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: IBM P.C. DOS 5.0
;; SOFTWARE: Word Perfect 5.1
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/312,157
;; FILING DATE: 14-May-1999
;; CLASSIFICATION: <Unknown>
;;
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/434,000
;; FILING DATE: <Unknown>
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Guise, Jeffrey W.
;; REGISTRATION NUMBER: 34,613
;; REFERENCE/DOCKET NUMBER: 212/127
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (619) 552-8400
;; TELEFAX: (619) 552-0159
;; TELEX: 67-351
;;
;; INFORMATION FOR SEQ ID NO: 6:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 757 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; DESCRIPTION: Bovine Polymunoglobulin Receptor
;;
;; US-09-312-157-6

Query Match 3.7%; Score 95; DB 4; Length 757;
Best Local Similarity 22.0%; Pred. No. 0.55;
Matches 88; Conservative 51; Mismatches 125; Indels 136; Gaps 26;

QY 4 KCG-----NCGPGYSTPLEAMKGPPE-----EYILPCLYNTGTEADYLATV 47
DB 109 KCGIGISSRGINFVSLVSDPAQASHAHYITDGLRTVINCPTFRANSERKSLC-- 166
-QY 48 DVDRSPQVC-QVTHR--LPMPLNKDELHHS--GMNTCSSCFGSDSTSRKLVLPSSLSS 102
DB 167 ---KKTIDCFQVVDSTGYGYSNKTDRAHISILSTNT-----LVFSVAIN 208
QY 103 RI-----YVVDGSEPRAPKIH--KVIEPK-----DIHAK-----CEIA----- 134
DB 209 RVKISDAGMYCCAGDADAKAKINIDQVLEPEBELVYGLDRSSVTFDCSLGPEVAVPK 268
QY 135 FLHTSHSLASEVMISLSGVKNGKGGFVL--DGETFEKKGWERG-----GAAP 185
DB 269 FLCCKKNGGACNVINILTKKAQDFGRIVSVPRDNGVFSVHTSLRKEDAGRYVCGAQP 328
QY 186 LG--YDFWYQPRHNMISTEWAP-----NVLRDG-----FNPADVEAGLYGSHLYV 230
DB 329 EGEFGDGMPIYQAWQLFVNEEALIPASPSYVKGAVGGSVTVSCPTNPADANSAKTWCH-- 385
QY 231 WDWORH---EIVOTLSLK-----DGLILEIRFLNPDAAQGFVGALSL 271
DB 386 WEEAONRCRCPLVESRGLMEQYEGRLVLLTEPENGTFTYVILNOLDODA--GFYWCV-- 441
QY 272 STIDRFKNKGGT-W--SVE-KVYQ-----VPPKKYKGM 302
DB 442 -----TDGDTKMTSTVELKVVQGEPSLKVPRKNVTAML 473

RESULT 15
US-08-856-841-16
; Sequence 16, Application US/08856841
; Patent No. 6274307
; GENERAL INFORMATION:
; APPLICANT: ERWIN SOUTSCHEK
; APPLICANT: MANFRED MOTZ
; TITLE OF INVENTION: IMMUNOLOGICALLY ACTIVE PEPTIDES
; TITLE OF INVENTION: OR POLYPEPTIDES FROM THE PARVOVIRUS B19
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BROOKS HAIDT HAFNER & DELAHUNTY
; STREET: 99 PARK AVENUE
; CITY: NEW YORK
; STATE: NY
; COUNTRY: USA
; ZIP: 10016
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" FLOPPY DISC
; COMPUTER: AT&T - IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS Version 6.2
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/856,841
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,658
; FILING DATE: 16-MARCH-1994
; APPLICATION NUMBER: US 07/917,096
; FILING DATE: 4-AUGUST-1992
; APPLICATION NUMBER: PCT/DE91/00106
; FILING DATE: 8-FEBRUARY-1991
; APPLICATION NUMBER: DE4038262
; FILING DATE: 8-FEBRUARY-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINSON, WILLIAM R.
; REGISTRATION NUMBER: 27,224
; REFERENCE/DOCKET NUMBER: LKR-9222-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 697-3355
; TELEFAX: (212) 557-5635
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 500
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PEPTIDE
; DESCRIPTION: N/A
; HYPOTHETICAL: N/A
; ANTI-SENSE: N/A
; FRAGMENT TYPE: INTERNAL
; ORIGINAL SOURCE: SERUM FROM PATIENT WITH ACUTE
; ORIGINAL SOURCE: INFECTION (ERYTHEMA INFECTIONOSUM)
; IMMEDIATE SOURCE: GENETICALLY ENGINEERED PEPTIDE
; POSITION IN GENOME: N/A
; FEATURE:
; NAME/KEY: N/A
; LOCATION: N/A
; IDENTIFICATION METHOD: amino acid analysis and
; IDENTIFICATION METHOD: mass spectrometry
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS: COSSART, Y.E.
; AUTHORS: FIELD, A.M.
; AUTHORS: CANT, B.
; AUTHORS: WIDDOWS, D.
; TITLE: PARVOVIRUS-LIKE PARTICLES IN HUMAN SERA
; JOURNAL: LANCET
; VOLUME: I
; ISSUE:

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:45:56 ; Search time 111.09 seconds

(without alignments)
1495.498 Million cell updates/sec

Title: US-09-841-758-1

Perfect score: 2550
Sequence: 1 MATKCGNCGPYSTPLEAMK.....PALAHLRYPGDCSSDIWI 472

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCYUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2550	100.0	472	22	US-09-841-758-1
2	2550	100.0	472	23	US-09-841-758-1
3	2550	100.0	472	26	US-60-239-841-11
4	2539	99.6	499	1	PCT-US01-14827-8378
5	2539	99.6	499	23	US-09-825-300-1717
6	2526	99.1	514	26	US-60-340-187-420
7	2435	95.5	472	22	US-09-841-758-3

8	2280	89.4	472	22	US-09-841-758-5
9	2240	87.8	472	22	US-09-841-758-4
10	1740.5	68.3	582	1	PCT-US01-08656-6715
11	1740.5	68.3	582	1	PCT-US01-14827-8378
12	1601.5	62.8	480	26	US-60-324-109-20719
13	1583.5	62.1	493	26	US-60-324-109-20454
14	1582.5	62.1	484	22	US-09-855-768-741
15	1571.5	61.6	490	21	US-09-708-427-21262
16	1562.5	61.3	488	20	US-09-620-394B-6447
17	1562.5	61.3	503	20	US-09-620-394B-6446
18	1545.5	60.6	487	20	US-09-620-394B-814
19	1545.5	60.6	502	20	US-09-620-394B-813
20	1523.5	59.7	455	21	US-09-708-427-21263
21	1514.5	59.4	455	20	US-09-620-394B-6448
22	1494.5	58.6	455	20	US-09-620-394B-815
23	1479	58.0	478	21	US-09-708-427-11627
24	1428	56.0	446	21	US-09-708-427-11628
25	1359.5	53.3	409	21	US-09-708-427-21264
26	1273	49.9	400	21	US-09-708-427-11629
27	1247.5	48.9	468	20	US-09-614-150-41685
28	1247.5	48.9	468	26	US-60-167-324-1108
29	1247.5	48.9	468	26	US-60-173-386-1072
30	1247.5	48.9	468	26	US-60-175-871-1219
31	1247.5	48.9	468	26	US-60-184-775-1103
32	1247.5	48.9	468	26	US-60-191-637-41319
33	1247.5	48.9	468	26	US-60-191-700-1185
34	1073.5	42.1	328	20	US-09-620-394B-5715
35	913.5	35.8	275	20	US-09-620-394B-8284
36	905.5	35.5	275	20	US-09-620-394B-5716
37	807	31.6	236	20	US-09-620-394B-8285
38	799	31.3	323	19	US-09-620-394B-5717
39	755.5	29.6	323	19	US-09-595-329A-662
40	755.5	29.6	325	19	US-09-595-329A-661
41	755.5	29.6	339	19	US-09-595-329A-660
42	713	28.0	199	20	US-09-620-394B-8286
43	549.5	21.5	135	20	US-09-621-976-6846
44	549.5	21.5	135	26	US-60-147-499-6846
45	542	21.3	229	16	US-09-270-767-44853

ALIGNMENTS

RESULT 1

US-09-841-758-1

Sequence 1 Application US/09841758

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/841,758

FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/088,641

FILING DATE: 1998-06-02

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0163 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 472 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: None

IMMEDIATE SOURCE:

LIBRARY: <unknown>

CLONE: 989953

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-841-758-1

Query Match 100.0%; Score 2550; DB 22; Length 472;
 Best Local Similarity 100.0%; Pred. No. 4.1e-252;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGCGYSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
 |||||||
 DB 1 MATKCGNGCGYSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
 |||||||
 QY 61 HRLPMPNLKDELHSGWNTCSSCFGDSSTKSRKLVLPSSLISRIYVVDGSEPRAPKLR 120
 |||||||
 DB 61 HRLPMPNLKDELHSGWNTCSSCFGDSSTKSRKLVLPSSLISRIYVVDGSEPRAPKLR 120
 |||||||
 QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSLGDYKGNKGGFVLLDGETFEVKGWERP 180
 |||||||
 DB 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSLGDYKGNKGGFVLLDGETFEVKGWERP 180
 |||||||
 QY 181 GGAAPLGDFWYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEIQ 240
 |||||||
 DB 181 GGAAPLGDFWYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEIQ 240
 |||||||
 QY 241 TSLKDGILPLEIRFLHNDDAAGFVGALSTIOREYKNEGTSVEKVIQVPPKRYG 300
 |||||||
 DB 241 TSLKDGILPLEIRFLHNDDAAGFVGALSTIOREYKNEGTSVEKVIQVPPKRYG 300
 |||||||
 QY 301 WLLPEMGLITDILLSLDRFLYFSNMHLGDLROYDISPQRPRLTGOLFSGSYKGP 360
 |||||||
 DB 301 WLLPEMGLITDILLSLDRFLYFSNMHLGDLROYDISPQRPRLTGOLFSGSYKGP 360
 |||||||
 QY 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMIQLSLDGKRLYITTSLSYAMDKQFYDPLIRE 420
 |||||||
 DB 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMIQLSLDGKRLYITTSLSYAMDKQFYDPLIRE 420
 |||||||
 QY 421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLGALAHLELRYGDCSSDIWI 472
 |||||||
 DB 421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLGALAHLELRYGDCSSDIWI 472
 |||||||

RESULT 2

US-09-981-353-11

Sequence 11, Application US/09981353

GENERAL INFORMATION:

APPLICANT: Lassek, Amy W.

TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER

FILE REFERENCE: PA-0038 US

CURRENT APPLICATION NUMBER: US/09/981,353

CURRENT FILING DATE: 2001-10-11

NUMBER OF SEQ ID NOS: 194

SOFTWARE: PERL Program

SEQ ID NO 11

LENGTH: 472

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc-feature

OTHER INFORMATION: Incyte ID No: 611514CD1

US-09-981-353-11

Query Match 100.0%; Score 2550; DB 23; Length 472;
 Best Local Similarity 100.0%; Pred. No. 4.1e-252;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGCGYSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
 |||||||
 DB 1 MATKCGNGCGYSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
 |||||||
 QY 61 HRLPMPNLKDELHSGWNTCSSCFGDSSTKSRKLVLPSSLISRIYVVDGSEPRAPKLR 120
 |||||||
 DB 61 HRLPMPNLKDELHSGWNTCSSCFGDSSTKSRKLVLPSSLISRIYVVDGSEPRAPKLR 120
 |||||||
 QY 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSLGDYKGNKGGFVLLDGETFEVKGWERP 180
 |||||||
 DB 121 VIEPKDIAKCELAFLHTSHCLASGEVMTSLGDYKGNKGGFVLLDGETFEVKGWERP 180
 |||||||
 QY 181 GGAAPLGDFWYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEIQ 240
 |||||||
 DB 181 GGAAPLGDFWYQPRHNWMISTEMAAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEIQ 240
 |||||||
 QY 241 TSLKDGILPLEIRFLHNDDAAGFVGALSTIOREYKNEGTSVEKVIQVPPKRYG 300
 |||||||
 DB 241 TSLKDGILPLEIRFLHNDDAAGFVGALSTIOREYKNEGTSVEKVIQVPPKRYG 300
 |||||||
 QY 301 WLLPEMGLITDILLSLDRFLYFSNMHLGDLROYDISPQRPRLTGOLFSGSYKGP 360
 |||||||
 DB 301 WLLPEMGLITDILLSLDRFLYFSNMHLGDLROYDISPQRPRLTGOLFSGSYKGP 360
 |||||||
 QY 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMIQLSLDGKRLYITTSLSYAMDKQFYDPLIRE 420
 |||||||
 DB 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMIQLSLDGKRLYITTSLSYAMDKQFYDPLIRE 420
 |||||||
 QY 421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLGALAHLELRYGDCSSDIWI 472
 |||||||
 DB 421 GSVMLQVDVDTYKGGIKLNPFLVDFGKEPLGALAHLELRYGDCSSDIWI 472
 |||||||

RESULT 3

US-60-239-841-11

Sequence 11, Application US/60239841

GENERAL INFORMATION:

APPLICANT: Lassek, Amy W.

TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER

FILE REFERENCE: PA-0038 P

CURRENT APPLICATION NUMBER: US/60/239,841

CURRENT FILING DATE: 2000-10-11

NUMBER OF SEQ ID NOS: 194

SOFTWARE: PERL Program

SEQ ID NO 11

LENGTH: 472

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc-feature

OTHER INFORMATION: Incyte ID No: 611514CD1

US-60-239-841-11

Query Match 100.0%; Score 2550; DB 26; Length 472;
 Best Local Similarity 100.0%; Pred. No. 4.1e-252;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATKCGNGCGYSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
 |||||||
 DB 1 MATKCGNGCGYSTPLEAMKGPREEIVYLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
 |||||||
 QY 61 HRLPMPNLKDELHSGWNTCSSCFGDSSTKSRKLVLPSSLISRIYVVDGSEPRAPKLR 120
 |||||||
 DB 61 HRLPMPNLKDELHSGWNTCSSCFGDSSTKSRKLVLPSSLISRIYVVDGSEPRAPKLR 120
 |||||||

```

QY 121 VIEPKDIHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGWTERP 180
    |||||||
Db 121 VIEPKDIHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGWTERP 180
QY 181 GGAAPLGIDFWYQPRHNVMISTEMAAPNVLRDGFNPADVAGLYGSHLYWDMQRHEIYQ 240
    |||||||
Db 181 GGAAPLGIDFWYQPRHNVMISTEMAAPNVLRDGFNPADVAGLYGSHLYWDMQRHEIYQ 240
QY 241 TSLKKGILPLEIRFLHNPAAAGFYGCALSSITQRFYKNEGTSWEKYIQVPPKKVKG 300
    |||||||
Db 241 TSLKKGILPLEIRFLHNPAAAGFYGCALSSITQRFYKNEGTSWEKYIQVPPKKVKG 300
QY 301 WLLPEMGLITDILLSDRFLYFSNMLHGDLEQYDISDPQRPRLTGOLFSGSIYKGP 360
    |||||||
Db 301 WLLPEMGLITDILLSDRFLYFSNMLHGDLEQYDISDPQRPRLTGOLFSGSIYKGP 360
QY 361 VOYLEDEELKSQPEPLVYVKGKRVAGGQPMQLSLDGKRLYITTSLSYAMDQKQYPLIRE 420
    |||||||
Db 361 VOYLEDEELKSQPEPLVYVKGKRVAGGQPMQLSLDGKRLYITTSLSYAMDQKQYPLIRE 420
QY 421 GSYMLQYDVDTYVKGKLNPNFLVDGKEPLGALAHLELRYPGDCSSDIWI 472
    |||||||
Db 421 GSYMLQYDVDTYVKGKLNPNFLVDGKEPLGALAHLELRYPGDCSSDIWI 472

```

RESULT 4

```

; Sequence 1717, Application PC/TUS0005988
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Human Prostate Cancer Associated Gene Sequences and Polypeptides
; FILE REFERENCE: PA101PCT
; CURRENT APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; EARLIER APPLICATION NUMBER: 60/124, 270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1717
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (486)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US00-05988-1717

```

```

Query Match          99.6%; Score 2539; DB 1; Length 499;
Best Local Similarity 99.6%; Pred. No. 6e-251;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MATKCGNCGYSTPLEAMGPREEIYLLPCIRNMGTEADYDLATVDDPKSPQCYVI 60
    |||||||
Db 28 MATKCGNCGYSTPLEAMGPREEIYLLPCIRNMGTEADYDLATVDDPKSPQCYVI 87
    |||||||
QY 61 HRLPMPNLKDELHSGMNTSCSCFGDSTKSRKTLVLPSSLISRIYVVDVSGSEPRAPKLK 120
    |||||||
Db 88 HRLPMPNLKDELHSGMNTSCSCFGDSTKSRKTLVLPSSLISRIYVVDVSGSEPRAPKLK 147
    |||||||
QY 121 VIEPKDIHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGWTERP 180
    |||||||

```

```

Db 148 VIEPKDIHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGWTERP 207
    |||||||
QY 181 GGAAPLGIDFWYQPRHNVMISTEMAAPNVLRDGFNPADVAGLYGSHLYWDMQRHEIYQ 240
    |||||||
Db 208 GGAAPLGIDFWYQPRHNVMISTEMAAPNVLRDGFNPADVAGLYGSHLYWDMQRHEIYQ 267
    |||||||
QY 241 TSLKKGILPLEIRFLHNPAAAGFYGCALSSITQRFYKNEGTSWEKYIQVPPKKVKG 300
    |||||||
Db 268 TSLKKGILPLEIRFLHNPAAAGFYGCALSSITQRFYKNEGTSWEKYIQVPPKKVKG 327
    |||||||
QY 301 WLLPEMGLITDILLSDRFLYFSNMLHGDLEQYDISDPQRPRLTGOLFSGSIYKGP 360
    |||||||
Db 328 WLLPEMGLITDILLSDRFLYFSNMLHGDLEQYDISDPQRPRLTGOLFSGSIYKGP 387
    |||||||
QY 361 VOYLEDEELKSQPEPLVYVKGKRVAGGQPMQLSLDGKRLYITTSLSYAMDQKQYPLIRE 420
    |||||||
Db 388 VOYLEDEELKSQPEPLVYVKGKRVAGGQPMQLSLDGKRLYITTSLSYAMDQKQYPLIRE 447
    |||||||
QY 421 GSYMLQYDVDTYVKGKLNPNFLVDGKEPLGALAHLELRYPGDCSSDIWI 472
    |||||||
Db 448 GSYMLQYDVDTYVKGKLNPNFLVDGKEPLGALAHLELRYPGDCSSDIWI 499
    |||||||

```

RESULT 5

```

; Sequence 1717, Application US/0925300
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124, 270
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1717
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (11)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (485)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (486)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1717

```

```

Query Match          99.6%; Score 2539; DB 23; Length 499;
Best Local Similarity 99.6%; Pred. No. 6e-251;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 MATKCGNCGYSTPLEAMGPREEIYLLPCIRNMGTEADYDLATVDDPKSPQCYVI 60
    |||||||
Db 28 MATKCGNCGYSTPLEAMGPREEIYLLPCIRNMGTEADYDLATVDDPKSPQCYVI 87
    |||||||
QY 61 HRLPMPNLKDELHSGMNTSCSCFGDSTKSRKTLVLPSSLISRIYVVDVSGSEPRAPKLK 120
    |||||||
Db 88 HRLPMPNLKDELHSGMNTSCSCFGDSTKSRKTLVLPSSLISRIYVVDVSGSEPRAPKLK 147
    |||||||
QY 121 VIEPKDIHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGWTERP 180
    |||||||
Db 148 VIEPKDIHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGFVLLDGETFEVKGWTERP 207
    |||||||
QY 181 GGAAPLGIDFWYQPRHNVMISTEMAAPNVLRDGFNPADVAGLYGSHLYWDMQRHEIYQ 240
    |||||||

```

```
Db 208 GGAAPLGDFWQPRHNWISTEMAPNVLKRGFNPADVEAGLYGSHLYVMDQREHIVQ 267
Oy 241 TLTKDGLPLEIRFLHNDAAQGFVGCALSTIQRFYKNEGTSVEKVIQVPPKKYKG 300
Db 268 TLTKDGLPLEIRFLHNDAAQGFVGCALSTIQRFYKNEGTSVEKVIQVPPKKYKG 327
Oy 301 WLPEMPGLITDILSLDDRFLYFSNMHLGDLROYDISDPORRLTGOLFGLGSIYKGGP 360
Db 328 WLPEMPGLITDILSLDDRFLYFSNMHLGDLROYDISDPORRLTGOLFGLGSIYKGGP 387
Oy 361 VOVLDEELKSQPEPLVYKGRVAGGPMIQLSLDGKRLYITTSLYSAMDKOFPDLIRE 420
Db 368 VOVLDEELKSQPEPLVYKGRVAGGPMIQLSLDGKRLYITTSLYSAMDKOFPDLIRE 447
Oy 421 GSVMLQVVDVTVKGGKLKLPNLFVDFGKEPLGPALAHLELRYGGCCSSDIWI 472
Db 448 GSVMLQVVDVTVKGGKLKLPNLFVDFGKEPLGPALAHXKRYGGCCSSDIWI 499
```

RESULT 6

US-60-340-187-420

Sequence 420, Application US/60340187

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Wang, Zhiwei

APPLICANT: Wang, Jian-Rui

APPLICANT: Ghosh, Malabika

APPLICANT: Meng, Gezh

APPLICANT: Boyle, Bryan J

APPLICANT: Dimanac, Radoje T

TITLE OF INVENTION: Novel Nucleic Acids and

FILE REFERENCE: Polypeptides

CURRENT FILING DATE: 2001-12-12

PRIOR APPLICATION NUMBER: US/60/340,187

PRIOR FILING DATE: 2000-01-21

PRIOR APPLICATION NUMBER: US 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: PCT/US00/35017

PRIOR FILING DATE: 2000-12-22

PRIOR APPLICATION NUMBER: US 09/491,404

PRIOR FILING DATE: 2000-01-25

PRIOR APPLICATION NUMBER: PCT/US01/02623

PRIOR FILING DATE: 2001-01-25

PRIOR APPLICATION NUMBER: US 09/496,914

PRIOR FILING DATE: 2000-02-03

PRIOR APPLICATION NUMBER: US 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: PCT/US01/03800

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: US 09/515,126

PRIOR FILING DATE: 2000-02-28

PRIOR APPLICATION NUMBER: US 09/577,409

PRIOR FILING DATE: 2000-05-18

PRIOR APPLICATION NUMBER: PCT/US01/04927

PRIOR FILING DATE: 2001-02-26

PRIOR APPLICATION NUMBER: US 09/519,705

PRIOR FILING DATE: 2000-03-07

PRIOR APPLICATION NUMBER: US 09/574,454

PRIOR FILING DATE: 2000-05-19

PRIOR APPLICATION NUMBER: PCT/US01/04941

PRIOR FILING DATE: 2001-03-05

PRIOR APPLICATION NUMBER: US 09/540,217

PRIOR FILING DATE: 2000-03-31

PRIOR APPLICATION NUMBER: US 09/649,167

PRIOR FILING DATE: 2000-08-23

PRIOR APPLICATION NUMBER: PCT/US01/08631

PRIOR FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: US 09/552,929

PRIOR FILING DATE: 2000-04-18

PRIOR APPLICATION NUMBER: US 09/770,160

```

; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US01/08656
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: PCT/US01/14827
; PRIOR FILING DATE: 2001-05-16
; NUMBER OF SEQ ID NOS: 1192
; SOFTWARE: P_Fl_genes Version 6.0
; SEQ ID NO 420
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-340-187-420
```

Query Match 99.1%, Score 2526; DB 26; Length 514;
Best Local Similarity 99.4%, Pred. No. 1,4e-249;
Matches 468; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```
Oy 2 ATKCGNCGPGYSTPLEAMKGPREEIVYLPCIYRNTGTEAPDYLATVDVDPKSPQCYIH 61
Db 44 ATKCGNCGPGYSTPLEAMKGPREEIVYLPCIYRNTGTEAPDYLATVDVDPKSPQCYIH 103
Oy 62 RLPEPNLKDLELHSGWNTCCSCFGDSTSRKTLVPLSISSRIYVVDGSEPRAKLHKV 121
Db 104 RLPEPNLKDLELHSGWNTCCSCFGDSTSRKTLVPLSISSRIYVVDGSEPRAKLHKV 163
Oy 122 IEPKDIHAKCELAFLHTSHCLASGEVMTSLGDVGNKGGFVLLDGETFEYKGTWERPG 181
Db 164 IEPKDIHAKCELAFLHTSHCLASGEVMTSLGDVGNKGGFVLLDGETFEYKGTWERPG 223
Oy 182 GAAPLGDFWQPRHNWISTEMAPNVLKRGFNPADVEAGLYGSHLYVMDQREHIVOT 241
Db 224 GAAPLGDFWQPRHNWISTEMAPNVLKRGFNPADVEAGLYGSHLYVMDQREHIVOT 283
Oy 242 LSLKGLPLPLEIRFLHNDAAQGFVGCALSTIQRFYKNEGTSVEKVIQVPPKKYKG 301
Db 284 LSLKGLPLPLEIRFLHNDAAQGFVGCALSTIQRFYKNEGTSVEKVIQVPPKKYKG 343
Oy 302 LLPEMPGLITDILSLDDRFLYFSNMHLGDLROYDISDPORRLTGOLFGLGSIYKGGPV 361
Db 344 LLPEMPGLITDILSLDDRFLYFSNMHLGDLROYDISDPORRLTGOLFGLGSIYKGGPV 403
Oy 362 QVLEDEELKSQPEPLVYKGRVAGGPMIQLSLDGKRLYITTSLYSAMDKOFPDLIREG 421
Db 404 QVLEDEELKSQPEPLVYKGRVAGGPMIQLSLDGKRLYITTSLYSAMDKOFPDLIREG 463
Oy 422 GSVMLQVVDVTVKGGKLKLPNLFVDFGKEPLGPALAHLELRYGGCCSSDIWI 472
Db 464 GSVMLQVVDVTVKGGKLKLPNLFVDFGKEPLGPALAHLELRYGGCCSSDIWI 514
```

RESULT 7

US-09-841-758-3

Sequence 3, Application US/09841758

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841.758
; FILING DATE: 24-Apr-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/088.641
; FILING DATE: 1998-06-02
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: None
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1374792
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-841-758-3

```

```

Query Match          95.5%; Score 2435; DB 22; Length 472;
Best Local Similarity 96.0%; Pred. No. 2.7e-240;
Matches 453; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

```

```

QY 1 MATCGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGEADYLATVVDPKSPQYCVI 60
D 1 MATCGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGEADYLATVVDPKSPQYCVI 60
QY 61 HRLPMPLKDELHSHGWTSCSCFGDSTKSRKTLVPSLISSRIYVVDVGSSEPRAPKLIK 120
D 61 HRLPMPLKDELHSHGWTSCSCFGDSTKSRKTLVPSLISSRIYVVDVGSSEPRAPKLIK 120
QY 121 VIEPRDIHAKCELAFHSHCLASGEVWISSLDGVKNGKGFVLLDDETFEVGWTWERP 180
D 121 VIEPRDIHAKCELAFHSHCLASGEVWISSLDGVKNGKGFVLLDDETFEVGWTWERP 180
QY 181 GGAAPLGYDFWYQPRHNVMYSTEMAAPNVLRDGENPADVEAGLYGSHLYVMDQRHETIQ 240
D 181 GGAAPLGYDFWYQPRHNVMYSTEMAAPNVLRDGENPADVEAGLYGSHLYVMDQRHETIQ 240
QY 241 TSLKDGILPLEIRFLNHPDAOGFVGCAVLSSTIQRETKNKGTSVEKVIQVPPKKVKG 300
D 241 TSLKDGILPLEIRFLNHPDAOGFVGCAVLSSTIQRETKNKGTSVEKVIQVPPKKVKG 300
QY 301 WLLPEMPEGLITDILSLDDRFLYFSNMWHLGDLROYDISDPQRPRLTGOLFSGSIVKGGP 360
D 301 WLLPEMPEGLITDILSLDDRFLYFSNMWHLGDLROYDISDPQRPRLTGOLFSGSIVKGGP 360
QY 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMQIOLSLDGKRLYTTTSLYSAMDQKQFYDLIRE 420
D 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMQIOLSLDGKRLYTTTSLYSAMDQKQFYDLIRE 420
QY 421 GSVMLQVDVDTVKGGLKLNPNFLVDFGKEPLGPAALAHLELRYPGGDCSSDIWI 472
D 421 GSVMLQVDVDTVKGGLKLNPNFLVDFGKEPLGPAALAHLELRYPGGDCSSDIWI 472

```

```

RESULT 8
; Sequence 5, Application US/09841758
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL HUMAN SELENIUM-BINDING PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
;

```

```

;
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: IBM Compatible
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/841.758
; FILING DATE: 24-Apr-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/088.641
; FILING DATE: 1998-06-02
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0163 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 298710
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-841-758-5

```

```

Query Match          89.4%; Score 2280; DB 22; Length 472;
Best Local Similarity 87.7%; Pred. No. 2.2e-224;
Matches 414; Conservative 27; Mismatches 31; Indels 0; Gaps 0;

```

```

QY 1 MATCGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGEADYLATVVDPKSPQYCVI 60
D 1 MATCGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGEADYLATVVDPKSPQYCVI 60
QY 61 HRLPMPLKDELHSHGWTSCSCFGDSTKSRKTLVPSLISSRIYVVDVGSSEPRAPKLIK 120
D 61 HRLPMPLKDELHSHGWTSCSCFGDSTKSRKTLVPSLISSRIYVVDVGSSEPRAPKLIK 120
QY 121 VIEPRDIHAKCELAFHSHCLASGEVWISSLDGVKNGKGFVLLDDETFEVGWTWERP 180
D 121 VIEPRDIHAKCELAFHSHCLASGEVWISSLDGVKNGKGFVLLDDETFEVGWTWERP 180
QY 181 GGAAPLGYDFWYQPRHNVMYSTEMAAPNVLRDGENPADVEAGLYGSHLYVMDQRHETIQ 240
D 181 GGAAPLGYDFWYQPRHNVMYSTEMAAPNVLRDGENPADVEAGLYGSHLYVMDQRHETIQ 240
QY 241 TSLKDGILPLEIRFLNHPDAOGFVGCAVLSSTIQRETKNKGTSVEKVIQVPPKKVKG 300
D 241 TSLKDGILPLEIRFLNHPDAOGFVGCAVLSSTIQRETKNKGTSVEKVIQVPPKKVKG 300
QY 301 WLLPEMPEGLITDILSLDDRFLYFSNMWHLGDLROYDISDPQRPRLTGOLFSGSIVKGGP 360
D 301 WLLPEMPEGLITDILSLDDRFLYFSNMWHLGDLROYDISDPQRPRLTGOLFSGSIVKGGP 360
QY 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMQIOLSLDGKRLYTTTSLYSAMDQKQFYDLIRE 420
D 361 VOVEDEELKSQPEPLVYVKGKRVAGGPMQIOLSLDGKRLYTTTSLYSAMDQKQFYDLIRE 420
QY 421 GSVMLQVDVDTVKGGLKLNPNFLVDFGKEPLGPAALAHLELRYPGGDCSSDIWI 472
D 421 GSVMLQVDVDTVKGGLKLNPNFLVDFGKEPLGPAALAHLELRYPGGDCSSDIWI 472

```

RESULT 9

US-09-841-758-4

Sequence 4, Application US/09841758

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

Hawkins, Phillip R.

TITLE OF INVENTION: NOVEL HUMAN SILENIUM-BINDING PROTEIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: US

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/841,758

FILING DATE: 24-Apr-2001

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/088,641

FILING DATE: 1998-06-02

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0163 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 472 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 227630

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-841-758-4

Query Match 87.8%; Score 2240; DB 22; Length 472;

Best Local Similarity 86.0%; Pred. No. 2.9e-220;

Matches 406; Conservative 29; Mismatches 37; Indels 0; Gaps 0;

QY 1 MATKCGNGPGYSTPLEAKKGPREEIYVLPCTYRNTGTAPDYLATVDVDPKSPQYCY 60
DB 1 MATKCKCKCPGSTPLEAKKGPREEIYVLPCTYRNTGTAPDYLATVDVDPKSPQYCY 60
QY 61 HRLPMPNLKDELHSHGWNCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPK 120
DB 61 HRLPMPNLKDELHSHGWNCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPK 120
QY 61 HRLPMPNLKDELHSHGWNCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPK 120
DB 61 HRLPMPNLKDELHSHGWNCCSCFGDSTKSRKLVLPSSLISRIYVVDGSEPRAPK 120
QY 121 VIEPKDIAHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGGFVLLDGETFEYK 180
DB 121 VIEPKDIAHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGGFVLLDGETFEYK 180
QY 121 VIEPKDIAHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGGFVLLDGETFEYK 180
DB 121 VIEPKDIAHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGGFVLLDGETFEYK 180
QY 181 GGAAPGAYGFYQPRHNVMISTEAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEI 240
DB 181 GGAAPGAYGFYQPRHNVMISTEAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEI 240
QY 181 GGAAPGAYGFYQPRHNVMISTEAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEI 240
DB 181 GGAAPGAYGFYQPRHNVMISTEAPNVLRDGFNPADVEAGLYGSHLYVMDMORHEI 240
QY 241 TSLKDGILPLEIRFLHNDAAAGFVGALSTTIORFYKNEGTVSVEKVIQVPRKVG 300
DB 241 TSLKDGILPLEIRFLHNDAAAGFVGALSTTIORFYKNEGTVSVEKVIQVPRKVG 300
QY 241 TSLKDGILPLEIRFLHNDAAAGFVGALSTTIORFYKNEGTVSVEKVIQVPRKVG 300
DB 241 TSLKDGILPLEIRFLHNDAAAGFVGALSTTIORFYKNEGTVSVEKVIQVPRKVG 300

QY 301 WLPEMPGLITDILSLDRLFLFSNWLHGLROYDISDPQRRLTGLFLGGSIVKGP 360
DB 301 WLPEMPGLITDILSLDRLFLFSNWLHGLROYDISDPQRRLTGLFLGGSIVKGP 360
QY 361 VOYLEDDELKQPEPLVYKGRVAGGPGMIOISLDGKRLYITTSLYSAMDQKQFYD 420
DB 361 VOYLEDDELKQPEPLVYKGRVAGGPGMIOISLDGKRLYITTSLYSAMDQKQFYD 420
QY 421 GSVMLOVDVYKGGKLNPNFLVDFGKEPLPALAHLELYPGDCSSDIWI 472
DB 421 GSVMLOVDVYKGGKLNPNFLVDFGKEPLPALAHLELYPGDCSSDIWI 472

RESULT 10

PCT-US01-08656-6715

Sequence 6715, Application PC/TUS0108656

GENERAL INFORMATION:

APPLICANT: Hysq, Inc

TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES

FILE REFERENCE: 21272-066

CURRENT APPLICATION NUMBER: PCT/US01/08656

CURRENT FILING DATE: 2001-04-16

PRIOR APPLICATION NUMBER: 09/522,929

PRIOR FILING DATE: 2000-04-18

PRIOR APPLICATION NUMBER: 09/770,160

PRIOR FILING DATE: 2001-01-26

NUMBER OF SEQ ID NOS: 10994

SOFTWARE: Custom

SEQ ID NO: 6715

LENGTH: 582

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)...(582)

OTHER INFORMATION: Xaa = X or * as defined in Table 2

PCT-US01-08656-6715

Query Match 68.3%; Score 1740.5; DB 1; Length 582;

Best Local Similarity 73.5%; Pred. No. 8.5e-169;

Matches 367; Conservative 21; Mismatches 78; Indels 33; Gaps 20;

QY 2 ATKCGNGPGYSTPLEAKKGPREEIYVLPCTYRNTGTAPDYLATVDVDPKSPQYCY 59
DB 37 ATKCGNGPGYSTPLEAKKGPREEIYVLPCTYRNTGTAPDYLATVDVDPKSPQYCY 96
QY 60 IRLPMPNLKDELHSHGWNCCSCFGDSTKSR-TKLVLPSSLISRIYVVDGSEPRAP 116
DB 97 IRLPMPNLKDELHSHGWNCCSCFGDSTKSR-TKLVLPSSLISRIYVVDGSEPRAP 156
QY 117 KLHK-VIEPKDIAHAKCELAFLHTSHCLASGEVMISSLDGVKNGKGGFVLLDGETFEYK 175
DB 157 KLAQSLSPRDIHAKCNMAFLHTSHCLASGEVMISSLDGVKNGKGGFVLLDGETFEYK 216
QY 176 TWEREGGAAPGAYGFYQPRHNVMISTEAPNVLRDGFNPADVEAGLYGSHLYVMDMOR 235
DB 217 TWEREGGAAPGAYGFYQPRHNVMISTEAPNVLRDGFNPADVEAGLYGSHLYVMDMOR 276
QY 236 HEIYVTLKDGILPLEIRFLHNDAAAGFVGALSTTIORFYKNEGTVSVEKVIQVPRK 291
DB 277 HEIYVTLKDGILPLEIRFLHNDAAAGFVGALSTTIORFYKNEGTVSVEKVIQVPRK 336
QY 292 QVPRKVGKMLPEMPGLITDILSLDRLFLFSNWLHGLROYDISDPQRRLTGLFLG 349
DB 337 QVPRKVGKMLPEMPGLITDILSLDRLFLFSNWLHGLROYDISDPQRRLTGLFLG 396
QY 350 FLGGSIVAGGPGVYLEDDELKQPEPLVYKGRVAGGPGMIOISLDGKRLYITTSLY 403
DB 397 FLGGSIVAGGPGVYLEDDELKQPEPLVYKGRVAGGPGMIOISLDGKRLYITTSLY 456
QY 404 SLYSAMDQKQFYD-----LIRGGSVMLOVDVYKGG-TKLNPNF-LVDFGK-EPLGP 453
DB 404 SLYSAMDQKQFYD-----LIRGGSVMLOVDVYKGG-TKLNPNF-LVDFGK-EPLGP 453

GenCore version 4.5
(c) 1993 - 2000 Com

Search time 37.67 Seconds

3039.753 Million cell updates/sec

US-09-841-758-1
2550
1 MATKCGNCGPYS

BLOSUM62

751360 seqs, 242600617 residues

its satisfying chosen parameters: 751360

```
length: 0
length: 200000000000
```

Minimum	Match	0%
---------	-------	----

Listing first 45 summaries

Pending_Patents_AA_New: *

```

: /cgn2_6/ploadate/2/paa/pgt_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US06_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US07_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US08_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US09_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US10_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US11_NEW_COMB.pep.*
: /cgn2_6/ploadate/2/paa/US12_NEW_COMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	length			
1	1590	62.4	492	7	US-00-391-781-1043	Sequence 1043, A
2	1587.5	62.3	476	5	US-00-935-625-2951	Sequence 2951, A
3	1587.5	62.3	476	5	US-00-935-625-12243	Sequence 12243, A
4	1587.5	62.3	476	5	US-00-935-625-16884	Sequence 16884, A
5	1587.5	62.3	476	5	US-00-935-625-28652	Sequence 28652, A
6	1587.5	62.3	480	5	US-00-935-625-2950	Sequence 2950, A
7	1587.5	62.3	480	5	US-00-935-625-12242	Sequence 12242, A
8	1587.5	62.3	480	5	US-00-935-625-16883	Sequence 16883, A
9	1587.5	62.3	480	5	US-00-935-625-28651	Sequence 28651, A
10	1538.5	60.3	455	5	US-00-935-625-2952	Sequence 2952, A
11	1538.5	60.3	455	5	US-00-935-625-12244	Sequence 12244, A
12	1538.5	60.3	455	5	US-00-935-625-16885	Sequence 16885, A
13	1538.5	60.3	455	5	US-00-935-625-28653	Sequence 28653, A
14	1535	60.2	490	5	US-00-935-625-5882	Sequence 5882, A
15	1535	60.2	490	5	US-00-935-625-30789	Sequence 30789, A
16	1484	58.2	458	5	US-00-935-625-5883	Sequence 5883, A
17	1484	58.2	458	5	US-00-935-625-30790	Sequence 30790, A
18	1479	58.0	478	5	US-00-935-625-5069	Sequence 5069, A
19	1444.5	56.6	469	5	US-00-935-625-4907	Sequence 4907, A
20	1444.5	56.6	502	5	US-00-935-625-4906	Sequence 4906, A
21	1428	56.0	446	5	US-00-935-625-5070	Sequence 5070, A
22	1339.5	52.5	409	5	US-00-935-625-4908	Sequence 4908, A
23	1329	52.1	412	5	US-00-935-625-5884	Sequence 5884, A
24	1329	52.1	412	5	US-00-935-625-30791	Sequence 30791, A
25	1273	49.9	400	5	US-00-935-625-5071	Sequence 5071, A
26	1178	46.2	235	5	US-00-629-16594-18237	Sequence 18237, A

27	920.5	36.1	291.5	US-09-935-625-58029	Sequence 5809, A
28	920.5	36.1	291.5	US-09-935-625-58021	Sequence 30421, A
29	913.5	35.8	275.5	US-09-935-625-10393	Sequence 10393, A
30	913.5	35.8	275.5	US-09-935-625-10392	Sequence 10392, A
31	807	31.6	236.5	US-09-935-625-10394	Sequence 10394, A
32	807	31.6	236.5	US-09-935-625-10394	Sequence 10394, A
33	770.5	30.2	239.5	US-09-935-625-5810	Sequence 5810, A
34	770.5	30.2	239.5	US-09-935-625-5810	Sequence 5810, A
35	770.5	30.2	239.5	US-09-935-625-30422	Sequence 30422, A
36	665	26.1	200.5	US-09-935-625-5811	Sequence 5811, A
37	665	26.1	200.5	US-09-935-625-30423	Sequence 30423, A
38	616	24.2	167.5	US-09-935-625-9670	Sequence 9670, A
39	601	23.6	219.5	US-09-935-625-10395	Sequence 10395, A
40	448.5	17.6	163.5	US-09-935-625-11324	Sequence 11324, A
41	448.5	17.6	163.5	US-09-935-625-5671	Sequence 5671, A
42	366	14.4	487.5	US-09-791-537-55612	Sequence 55612, A
43	366	14.4	487.5	US-60-360-039-6062	Sequence 6062, A
44	108.5	4.3	3427.7	US-60-389-987-1624	Sequence 1624, A
45	108.5	4.1	893.5	US-09-791-537-24989	Sequence 24989, A
46	103.5	4.1	407.5	US-09-791-537-66766	Sequence 66766, A

ALIGNMENTS

```

RESULT      1
US-60-391-781-1043
; Sequence 1043, Application US/60391781
;
; GENERAL INFORMATION:
; APPLICANT: Laitrie, Cathy C.
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)B
; CURRENT APPLICATION NUMBER: US/60/391,781
; PRIORITY FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: 60/365,301
; PRIORITY FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 3034
;
; SEQ ID NO 1043
; LENGTH: 492
;
; TYPE: PRT
;
; ORGANISM: Zea mays
;
; US-60-391-781-1043

```

Query Match	62.4%;	Score 1590;	DB 7;	Length 492;
Best Local Similarity	61.9%;	Pred. No. 4.2e-132;		
Matches 289;	Conservative 67;	Mismatches 107;	Indels 4;	Gaps 4;

[illegible]

Db 386 DDGDEQUNYRQVKGHHLRGGFRQIQJSLDGKRIYIYNLSLFESKWBDEFFEDDYVKKGSIML 445

QY 426 QVDDVTVKSGIKLNPFLVDEGKEPGLRALAHETRYFGSDSDIWI 472
I::IIII IIII :IIII IIII II IIIII::IIIIII::IIII
446 QIVDTEKGGIAVYPNFFYVDEGTEPRDRALAHETRYFGSGTSDIWI 492

```

RESULT      2
US-09-935-625-2951
: Sequence 2951, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
: OF INVENTION: MODULATING VARIOUS RESPONSES
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 2951
: LENGTH: 476
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..476
: OTHER INFORMATION: Ceres Seq. ID no. 2175254
: US-09-935-625-2951

```

Query Match	62.3%	Score 1587.5	DB 5	Length 476
Best Local Similarity	61.3%	Pred. No. 6.7e-132		
Matches 288; Conservative	63	Mismatches 116	Indels 3	Gaps 3

[illegible]

```

RESULT      3
US-09-935-625-12243
; Sequence 12243, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES,
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; AND METHODS THEREOF CAPABLE

```

```

FILE REFERENCE: 2750-1481P
CURRENT APPLICATION NUMBER: US/09/935,625
CURRENT FILING DATE: 2001-08-24
NUMBER OF SEQ ID NOS: 33136
SEQ ID NO 12243
LENGTH: 476
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: peptide
LOCATION: 1..476
OTHER INFORMATION: Ceres Seq. ID no. 2175254
US-09-935-625-12243

```

Query Match	62.3%	Score 1587.5;	DB 5;	Length 476;
Best Local Similarity	61.3%	Pred. No. 6.7e-132;		
Matches 288;	Conservative 63;	Mismatches 116;	Indels 3;	Gaps 3;

Qy	5	CGNCGPGVSTLEAMKMGCREIVLPCIRYNTGTAEADYDLYATVDVDRKSQYCCVHRLP	64
Dd	8	CKSGPGATFLLAMSGREKLIIYVAIYTGTGAKDYLATVDEPSSSYSSVHRLP	67
Qy	65	MPNIDELHSHGAMWTCSSCFGDSFKSRKLYLPSLISRIYVVDYGPAPRAPIKHYEP	124
Dd	68	MPYLEDELHSHGAMWSSCSCYDSCSCERRYLLPILLSGRYIVDTKINPREPSLHKFVDP	127
Qy	125	KDIHAKCGLAFLHSHCLASGEWMISSLGDVKKGKGK - GPVLLDGEFEVKGEMERGGA	183
Dd	128	AEVLEKTEGLAAPHOPHCCLASSGDVLYSCIGDDGNAEBSGFLLDSE - FNIKGRNEKJGNS	186
Qy	184	APLGEFWYQCPRHNVMIISTEMAAPENVLRDGNPDAVEAGLYGSHLYWDMQRIEYOTLS	243
Dd	187	PLXGDEFWYQCPRHNTWISTMGAPAAFLKFGDDLKDVSDGLYKRLHLYSWPQGLKQILD	246
Qy	244	LKD - GLPDLRLHNDPAAGFYGCALSSITQFFYKNEGTSVXEKVIQYPRKKYKGYL	302
Dd	247	LGDGLPLLEVRFLHEPDKAIFGACALSTVLYRFFKNDDTWSHEVAISVEPLUKVNMII	306
Qy	303	LPEMFGITLTDILSLDPRIFYSNWMLHGDLEQYOISIPQPRRLGQGLFGSIIYKGGPVQ	362
Dd	307	LPEMFGITLTDILSLDPRIFYSNWMLHGDLEQYOISIPQPRRLGQGLFGSIIYKGGPVQ	366
Qy	363	VLEDEELKSQEPPLVYKGRVAGSPQMILQSLDGKRLYITTSLYSAMDKOPEPDLREGS	422
Dd	367	ALGEGCAKAFQDPVARIKIGQRLRGCPQMFQSLDCKRLYVTNSLFSVMDROFPELPYKGS	426
Qy	423	VMLQVQDVYTVKGLKINPNFLVDSGKEPLGALAHLEIRPBGDCSSIIWI	472
Dd	427	HMLQVQDVYTVKGLKINPNFLVDSGKEPLGALAHLEIRPBGDCSSIIWI	476

```

RESULT      4
US-09-935-625-16884
: Sequence 16884, Application US/09935625
: GENERAL INFORMATION:
: APPLICANT: N. ALEXANDROV et al.
: TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPAB
: TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
: FILE REFERENCE: 2750-1481P
: CURRENT APPLICATION NUMBER: US/09/935,625
: CURRENT FILING DATE: 2001-08-24
: NUMBER OF SEQ ID NOS: 33136
: SEQ ID NO 16884
: LENGTH: 476
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: peptide
: LOCATION: 1..476
: OTHER INFORMATION: Ceres Seq. ID no. 2175254
US-09-935-625-16884

```



```
; OTHER INFORMATION: Ceres Seq. ID no. 2175253
US-09-935-625-28651

Query Match      62.3%; Score 1587.5; DB 5; Length 480;
Best Local Similarity 61.3%; Pred. No. 6,7e-132;
Matches 288; Conservative 63; Mismatches 116; Indels 3; Gaps 3;

QY 5 CGNCGPGYSPFLAMKMPREIEIVLPCIRNTGTEAPDYLATVDVDPKSPQYCOVIRHLP 64
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 12 CCKGSGPETAIRPLAMSGPREKIIYAAIYTGQAKPDYLATVVEPSSSTYSYVIRHLP 71
QY 65 MPNKLDELHSGMWTSCSCFGDSFKSRKTLVPSLISSRIYVDVGSEPRAPKLHKYIEP 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 72 MPYLEDELHSGMNSCSCYGDSSCERRYLILPILLSGRIVYIDTKINPREPSLHKFYDP 131
QY 125 KDIAKCELAFIHTSHCLASGEVMISSLDYKNGKG-GEVLLDGEFPEVKGTERRPGA 183
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 132 AEVLEKTLGLAVPHQPHCLASGDLVLSCLGDEDAEAGSGFLLDSE-FNIGKRWKDGNS 190
QY 184 APLEGDFWYOPRHNWMTSTEMAAPNVLRDGFNPADVEAGLGSILYVMDQRIEIVOTLS 243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 191 PLYGDFWYOPRHKMTIMSTSGAPAAFTKGFDLKDVSDGLGKHLHYVSWPQGLKQILD 250
QY 244 LKD-GLIPLEIRFLHNDAAQFVGCALSSTIQRFYKNEGTSVEKYIYPPKKVKGWL 302
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 251 LGDTGLPLEYRFLHEDPKATGFAGCALSSTLVRFKNDETWSEHVAISVEPLKVENMI 310
QY 303 LPEMPGLITDILSLDRFLYFSNWLHGDLRQYDISDPORPLTGOLFSGSIYKGPVQ 362
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 311 LPEMPGLITFLISLDRFLYCSNWLHGDLRQYINIEDPKPVLLGQIHVGLVOKGSLYL 370
QY 363 VLEDEELKSQDEPLVVGKRVAGGPOMIOLSLDGKRLYITTSLSAMDKOFYPLIREGS 422
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 371 ALGEGCAFQPDYVKIKGQRLRGSPQMFQSLDGKRLYVNSLFSVMDROFYPELVEKGS 430
QY 423 VMLQVDVYVVGGLKLNPNFLVDFGKEPLGALAHLELRYPGDCSSDIWI 472
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 431 HMLQIDVDYDQKGLSINPNFVDFGTEPDGSLAHEMRYPGDCTSDIYW 480

RESULT 10
US-09-935-625-2952
; Sequence 2952, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 2952
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..455
; OTHER INFORMATION: Ceres Seq. ID no. 2175255
US-09-935-625-2952

Query Match      60.3%; Score 1538.5; DB 5; Length 455;
Best Local Similarity 61.2%; Pred. No. 1.4e-127;
Matches 279; Conservative 62; Mismatches 112; Indels 3; Gaps 3;

QY 19 MKGPREIIVLPICIRNTGTEAPDYLATVDVDPKSPQYCOVIRHLPMPNKLDELHSGMN 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MSGPREKIIYAAIYTGQAKPDYLATVVEPSSSTYSYVIRHLPMPYLEDELHSGMN 60
QY 79 TCSSCFDSTSKRTKLVPSLISSRIYVDVGSEPRAPKLHKYIEPDKIAKCELAFIHT 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 61 SCSSCYDSSCERRYLILPILLSGRIVYIDTKINPREPSLHKFYDPAVELEKTLGLAVPHQ 120
QY 139 SHCLASGEVMTSSLDYKNGKG-GEVLLDGEFPEVKGTERRPGAAPLYGDFWYOPRHN 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 PHCLASGDLVLSCLGDEDAEAGSGFLLDSE-FNIGKRWKDGNSPLXYGDFWYOPRHK 179
QY 198 VMISTEWAAPNVLRDGFNPADVEAGLGSILYVMDQRIEIVOTLSKD-GLIPLEIRFL 256
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 TMISTWGAPAAFTKGFDLKDVSDGLGKHLHYVSWPQGLKQILDGDTGLPLEYRFL 239
QY 257 HNPDAAGFVGCALSSTIQRFYKNEGTSVEKYIYPPKKVKGWLPEMPGLITDILLS 316
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 HEPDKATGFAGCALSSTLVRFKNDETWSEHVAISVEPLKVENMILPEMPGLITDILIS 299
QY 317 LDDRFLEFSNWLHGDLRQYDISDPORPLTGOLFSGSIYKGPVQVLEDEELKSQDEPL 376
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 300 LDDRFLEFSNWLHGDLRQYINIEDPKPVLLGQIHVGLVOKGSLYLALGEGCAFQDVP 359
QY 377 VVGKRVAGGPOMIOLSLDGKRLYITTSLSAMDKOFYPLIREGSVMLQVDVYVVGGL 436
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 360 KIKGQRLRGSPQMFQSLDGKRLYVNSLFSVMDROFYPELVEKGSMLQIDVDYDQKGL 419
QY 437 KLNPNFLVDFGKEPLGALAHLELRYPGDCSSDIWI 472
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 420 SINPNFVDFGTEPDGSLAHEMRYPGDCTSDIYW 455

RESULT 11
US-09-935-625-12244
; Sequence 12244, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPA
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 12244
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..455
; OTHER INFORMATION: Ceres Seq. ID no. 2175255
US-09-935-625-12244

Query Match      60.3%; Score 1538.5; DB 5; Length 455;
Best Local Similarity 61.2%; Pred. No. 1.4e-127;
Matches 279; Conservative 62; Mismatches 112; Indels 3; Gaps 3;

QY 19 MKGPREIIVLPICIRNTGTEAPDYLATVDVDPKSPQYCOVIRHLPMPNKLDELHSGMN 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MSGPREKIIYAAIYTGQAKPDYLATVVEPSSSTYSYVIRHLPMPYLEDELHSGMN 60
QY 79 TCSSCFDSTSKRTKLVPSLISSRIYVDVGSEPRAPKLHKYIEPDKIAKCELAFIHT 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 SCSSCYDSSCERRYLILPILLSGRIVYIDTKINPREPSLHKFYDPAVELEKTLGLAVPHQ 120
QY 139 SHCLASGEVMTSSLDYKNGKG-GEVLLDGEFPEVKGTERRPGAAPLYGDFWYOPRHN 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 PHCLASGDLVLSCLGDEDAEAGSGFLLDSE-FNIGKRWKDGNSPLXYGDFWYOPRHK 179
QY 198 VMISTEWAAPNVLRDGFNPADVEAGLGSILYVMDQRIEIVOTLSKD-GLIPLEIRFL 256
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 180 TMISTWGAPAAFTKGFDLKDVSDGLGKHLHYVSWPQGLKQILDGDTGLPLEYRFL 239
QY 257 HNPDAAGFVGCALSSTIQRFYKNEGTSVEKYIYPPKKVKGWLPEMPGLITDILLS 316
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 240 HEPDKATGFAGCALSSTLVRFKNDETWSEHVAISVEPLKVENMILPEMPGLITDILIS 299
```



```

; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..490
; OTHER INFORMATION: Ceres Seq. ID no. 3089923
US-09-935-625-5882

```

```

Query Match          60.2%; Score 1535; DB 5; Length 490;
Best Local Similarity 59.4%; Pred. No. 3,1e-127;
Matches 281; Conservative 68; Mismatches 118; Indels 6; Gaps 4;

```

```

QY 5 CGNGCGYSPLEAMKGPREEIVYLPCLYRNTGTEADYATVVDPKSPQYCOVIHRLP 64
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 19 CCKSGPGYAPPLAMAGPREKLIYVYALYSTGDKDPYATVVDPSPESSVIRHLK 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 MPNKLDELHSGWNTCCSCFGDSTKSRKLYLPBLLISSRIYVVDGSEPPAPKLHKYIEP 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 MPYIGDELHHTGWNWSCSSCHGDAADRRYLVLPBLLISGRITAYIDTKDPRAPSLKYVER 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 KDIAKCELAFLHTSHCLASGEVMISSLDGVKNGKG-GEVLLDGETFEVKGWERPGGA 183
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 139 KEIAEKTGLAFPHHTSHCLASGDMLVSCLDGKEGNAKNGFLLDSD-FNVKSRMDKRGHG 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 184 APLEGYDFWYQPRHNVMISTEMAAPNVLRDGFNPADVEAGLYSHLYWDMQRIEIVOTLS 243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 198 PLFGYDFWYQPRFKMTISTWGAPKAFSKGFNLQHVADGLYSHLHYQMEGEMKQIID 257
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 244 L-KXGLPLEIRFLHNPDAAGFVGCALSSRIOREYNEGTWS---VEKYIYPPPKVK 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 258 LGNTGLPLEIRFLHNDPSKDTGYGSAISSNMIRFFKNSDDTWSHEASEVVISVPLKVE 317
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 300 GWLLPEMPGLITDILLSLDRFLYFSNWLHGDLRQYDISPQRPRLTGOLFSGSIYKGG 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 318 NWILPEMPGLITDPLISLDRFEFFVNMHLGDIHQYNIEDPKNVLTGQIHWGGLQKGS 377
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 360 PVQVLEDEBELKSQPEPLVYKGRVAGGPOMIQLSLDGKRLYITTSLSAMDKOFYPLLIR 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 378 PYKAVGEDGNTYQPDVPOIKGKSLRAGPOMIQLSLDGKRLYATNSLFSAMDROFYPRIME 437
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 420 EGSVMLQVDVTVKGLKLPNPNFLVDGKEPLGALAHLELRYPGDCSDIWI 472
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 438 KGSHTIQIDVDTDKGGLTLNPFVVDGDEPDGPALAHLEMYPGDCTSDIWI 490
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 15
US-09-935-625-30789
; Sequence 30789, Application US/09935625
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES, POLYPEPTIDES, CELLS, AND METHODS THEREOF CAPABLE
; TITLE OF INVENTION: MODULATING VARIOUS RESPONSES
; FILE REFERENCE: 2750-1481P
; CURRENT APPLICATION NUMBER: US/09/935,625
; CURRENT FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 33136
; SEQ ID NO 30789
; LENGTH: 490
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: peptide
; LOCATION: 1..490
; OTHER INFORMATION: Ceres Seq. ID no. 3089923
US-09-935-625-30789

```

```

Query Match          60.2%; Score 1535; DB 5; Length 490;
Best Local Similarity 59.4%; Pred. No. 3,1e-127;
Matches 281; Conservative 68; Mismatches 118; Indels 6; Gaps 4;

```

```

QY 5 CGNGCGYSPLEAMKGPREEIVYLPCLYRNTGTEADYATVVDPKSPQYCOVIHRLP 64
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 19 CCKSGPGYAPPLAMAGPREKLIYVYALYSTGDKDPYATVVDPSPESSVIRHLK 78
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

QY 65 MPNKLDELHSGWNTCCSCFGDSTKSRKLYLPBLLISSRIYVVDGSEPPAPKLHKYIEP 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 MPYIGDELHHTGWNWSCSSCHGDAADRRYLVLPBLLISGRITAYIDTKDPRAPSLKYVER 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 125 KDIAKCELAFLHTSHCLASGEVMISSLDGVKNGKG-GEVLLDGETFEVKGWERPGGA 183
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 139 KEIAEKTGLAFPHHTSHCLASGDMLVSCLDGKEGNAKNGFLLDSD-FNVKSRMDKRGHG 197
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 184 APLEGYDFWYQPRHNVMISTEMAAPNVLRDGFNPADVEAGLYSHLYWDMQRIEIVOTLS 243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 198 PLFGYDFWYQPRFKMTISTWGAPKAFSKGFNLQHVADGLYSHLHYQMEGEMKQIID 257
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 244 L-KXGLPLEIRFLHNPDAAGFVGCALSSRIOREYNEGTWS---VEKYIYPPPKVK 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 258 LGNTGLPLEIRFLHNDPSKDTGYGSAISSNMIRFFKNSDDTWSHEASEVVISVPLKVE 317
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 300 GWLLPEMPGLITDILLSLDRFLYFSNWLHGDLRQYDISPQRPRLTGOLFSGSIYKGG 359
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 318 NWILPEMPGLITDPLISLDRFEFFVNMHLGDIHQYNIEDPKNVLTGQIHWGGLQKGS 377
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 360 PVQVLEDEBELKSQPEPLVYKGRVAGGPOMIQLSLDGKRLYITTSLSAMDKOFYPLLIR 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 378 PYKAVGEDGNTYQPDVPOIKGKSLRAGPOMIQLSLDGKRLYATNSLFSAMDROFYPRIME 437
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 420 EGSVMLQVDVTVKGLKLPNPNFLVDGKEPLGALAHLELRYPGDCSDIWI 472
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 438 KGSHTIQIDVDTDKGGLTLNPFVVDGDEPDGPALAHLEMYPGDCTSDIWI 490
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Search completed: August 29, 2002, 16:50:24
Job time: 248 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:45:21 ; Search time 20.9 seconds
(without alignments)
2170.056 Million cell updates/sec

Title: US-09-841-758-1

Perfect score: 2550

Sequence: 1 MATKCGNCGPSTPLFAMK.....PALAHLRYPGDCSSDIWI 472

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2435	85.5	472	2	G01872
2	2254	88.4	472	2	S27878
3	1571.5	61.6	490	2	D71401
4	1524	59.8	1070	2	E71401
5	1479	58.0	478	2	A85153
6	1347.5	52.8	471	2	T26612
7	901.5	35.4	466	2	A99464
8	685.5	26.9	576	2	T24193
9	366	14.4	487	1	T32635
10	111	4.4	885	1	S67660
11	109.5	4.3	764	2	B82303
12	105	4.1	893	2	F72253
13	103.5	4.1	407	1	A55515
14	102.5	4.0	720	2	E84456
15	102.5	4.0	781	1	VCPV19
16	101.5	4.0	558	2	A13394
17	100.5	3.9	418	2	AH0184
18	100.5	3.9	980	2	T05414
19	99.5	3.9	601	2	T42251
20	99	3.9	845	1	GLVK
21	98.5	3.9	530	1	TVTK
22	98	3.8	507	2	AF3620
23	97.5	3.8	566	2	S07508
24	97.5	3.8	1164	2	G82100
25	97	3.8	757	2	T45956
26	96.5	3.8	1396	2	T10627
27	95.5	3.7	2925	2	T00133
28	95	3.7	757	1	S48841
29	95	3.7	3066	1	JQ1661

30	94.5	3.7	1070	2	JC4593	protein-tyrosine k
31	94.5	3.7	2367	2	S70172	toxin B - Clostrid
32	94	3.7	833	2	G75621	arylesterase/monox
33	93.5	3.7	369	2	S68692	deoxyhypusine synt
34	93.5	3.7	619	2	A69352	conserved hypothet
35	93.5	3.7	944	2	T18911	hypothetical prote
36	93	3.6	356	2	E83416	hypothetical prote
37	93	3.6	603	2	G82738	gamma-glutamyltran
38	92.5	3.6	588	2	S10893	transforming prote
39	92.5	3.6	1374	2	AE3259	extracellular seri
40	91.5	3.6	456	2	D97662	virulence acvB pro
41	91.5	3.6	456	2	AG2886	agrobacterium chro
42	91.5	3.6	929	2	G72677	hypothetical prote
43	91.5	3.6	1426	2	A99580	hypothetical prote
44	91	3.6	280	2	T48244	hypothetical prote
45	91	3.6	784	2	PN0009	neurofilament trip

ALIGNMENTS

Query Match	Best Local Similarity	Score	DB 2;	Length	472;
Matches 453; Conservative	96.0%;	Pred. No. 8.8e-193;	Mismatches 16;	Indels	0; Gaps 0;
1	MATKCGNCGPSTPLFAMKPREIYVLPCTYNTGTEADPDYATVDVDPKSPQCOVI	60			
1	MATKCGNCGPSTPLFAMKPREIYVLPCTYNTGTEADPDYATVDVDPKSPQCOVI	60			
61	HLRPMNLKDELHSGNNTSCSGDSTKSRKLVLPSSLSSRIYVDVGSSEPRAPRLK	120			
61	HLRPMNLKDELHSGNNTSCSGDSTKSRKLVLPSSLSSRIYVDVGSSEPRAPRLK	120			
121	VIEKRDIAKELAFITSHCLASGEVYVSSLGKVGKGGFVLDDGETFEVGTWERP	180			
121	VIEKRDIAKELAFITSHCLASGEVYVSSLGKVGKGGFVLDDGETFEVGTWERP	180			
121	VIEKRDIAKELAFITSHCLASGEVYVSSLGKVGKGGFVLDDGETFEVGTWERP	180			
181	GGAAPLGVDEWYORRHNMTSTENAAAPVLRDGNPAVDVEAGLSHLVYDMQRHELVQ	240			
181	GGAAPLGVDEWYORRHNMTSTENAAAPVLRDGNPAVDVEAGLSHLVYDMQRHELVQ	240			
181	GGAAPLGVDEWYORRHNMTSTENAAAPVLRDGNPAVDVEAGLSHLVYDMQRHELVQ	240			
241	TLSLKDGILPLIEIFLNPDAAGFVGCALSTTQREYKNGGTVSEKVIQVPPKVKG	300			
241	TLSLKDGILPLIEIFLNPDAAGFVGCALSTTQREYKNGGTVSEKVIQVPPKVKG	300			
241	TLSLKDGILPLIEIFLNPDAAGFVGCALSTTQREYKNGGTVSEKVIQVPPKVKG	300			
301	WLPDMGELTDLISLDDRLYFSNMLHGLDROYDISDPORPRLTGOLFSGSIVAGGP	360			
301	WLPDMGELTDLISLDDRLYFSNMLHGLDROYDISDPORPRLTGOLFSGSIVAGGP	360			
301	WLPDMGELTDLISLDDRLYFSNMLHGLDROYDISDPORPRLTGOLFSGSIVAGGP	360			
361	VQVLEDELKQSPRLVYKGRVAGGPMOTSLDGRKLYTTSTLVSAMKQFPPDLIRE	420			
361	VQVLEDELKQSPRLVYKGRVAGGPMOTSLDGRKLYTTSTLVSAMKQFPPDLIRE	420			
361	VQVLEDELKQSPRLVYKGRVAGGPMOTSLDGRKLYTTSTLVSAMKQFPPDLIRE	420			
421	GSVNLQVVDVIVKGLKLNPNFLVDVFGKEPLGALAHLELRYPGDCSSDIWI	472			

```

Db 421 GSVMLQVDVDTVVGKGLKLPNCLVDFGKEPLGALAHRLRPGDCCSDIWI 472
|||||
RESULT 2
S27878
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Jun-2000
C:Accession: S27878; JN0271
R:Bansel, M.P.; Cook, R.; Oborn, C.J.; Scott, J.; Mukhopadhyay, T.; Medina, D.
Submitted to the EMBL Data Library, February 1991
A:Description: Molecular cloning and sequencing of a liver protein which binds selenium.
A:Reference number: S27878
A:Accession: S27878
A:Molecule type: mRNA
A:Residues: 1-472 <BAN>
A:Cross-references: EMBL:M2032; NID:g200951; PIDN:AAA0104.1; PID:g200952
R:Pumford, N.R.; Martin, B.W.; Hinson, J.A.
Biochem. Biophys. Res. Commun. 182, 1348-1355, 1992
A:Title: A metabolite of acetaminophen covalently binds to the 56 kDa selenium binding F
A:Reference number: JN0271; MUID:92171951
A:Accession: JN0271
A:Molecule type: Protein
A:Residues: 72-73; 'K'; 175-181; 'G', 183; 'S', 185-189; 196-220; 228-242; 290-296; 334-343; 399-40
C:Comment: This protein covalently binds a metabolite of acetaminophen. This binding con
C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5
C:Keywords: liver; selenium binding

Query Match 88.4%; Score 2254; DB 2; Length 472;
Best Local Similarity 86.4%; Pred. No. 7.5e-178;
Matches 408; Conservative 29; Mismatches 35; Indels 0; Gaps 0;

OY 1 MATCGNCGPSTPLPAMKGPREEIYVLPCTIRNTGTEADYATVVDPKSPQYCOVI 60
|||
Db 1 MATCTCKGPGSTPLPAMKGPREEIYVLPCTIRNTGTEADYATVVDPKSPQYCOVI 60
|||
OY 61 HRLPMPNLKDELHSHGWNCCSCFCGDSRTKSLVPLSLISSRIYVVDGSEPRAPKLHK 120
|||||
Db 61 HRLPMPYLKDELHSHGWNCCSCFCGDSRTKSLVPLSLISSRIYVVDGSEPRAPKLHK 120
|||||
OY 121 VIEKQDIHAKELAFHLSHCLASGEVWISSLDGVKNGKGFVLLDGETFEVGTWERP 180
|||||
Db 121 VIESEIQAKCNVSSLSHCLASGEVWISFLGDQNGKGSFVLLDGETFEVGTWERP 180
|||||
OY 181 GGAAPLIGDFWYQPRHNWMTSTEMAAPVLRDGENPADAVEAGLSHLYWDMQRIHEIVQ 240
|||||
Db 181 GDAAPMGDFWYQPRHNWMTSTEMAAPVLRDGENPADAVEAGLSHLYWDMQRIHEIVQ 240
|||||
OY 241 TLSLKDGILPILRFLNPDAAAGFCVCAALSTIQRFYKNEGTVSWVEKVIQVPPKKYKG 300
|||||
Db 241 TLMQMDGILPILRFLNPDAAAGFCVCAALSTIQRFYKNEGTVSWVEKVIQVPPKKYKG 300
|||||
OY 301 WLLPEMRELITDIILSLDDRFLYFSNMLHGLROYDISDPQRPRLTQGLFSGSYVKGCP 360
|||||
Db 301 WMLPGVPELITDIILSLDDRFLYFSNMLHGLROYDISDPQRPRLTQGLFSGSYVKGCS 360
|||||
OY 361 VOVLDEELKQPEPLVYVKGKRVAGPOMIQLSDGKRLTYTTSLSYAMDQKQFPDLIRE 420
|||||
Db 361 VOVLDEELKQPEPLVYVKGKRVAGPOMIQLSDGKRLTYTTSLSYAMDQKQFPDLIRE 420
|||||
OY 421 GSVMLQVDVDTVVGKGLKLPNCLVDFGKEPLGALAHRLRPGDCCSDIWI 472
|||||
Db 421 GSVMLQVDVDTVVGKGLKLPNCLVDFGKEPLGALAHRLRPGDCCSDIWI 472
|||||

RESULT 3
D71401
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000

```

```

C:Accession: D71401
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneve, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chaitatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113
A:Accession: D71401
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-490 <BEV>
A:Cross-references: GB:297335; NID:g2244747; PID:g2244759
C:Genetics:
A:Map position: ACOP9-4G3845
C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 61.6%; Score 1571.5; DB 2; Length 490;
Best Local Similarity 61.1%; Pred. No. 1.5e-121;
Matches 287; Conservative 65; Mismatches 115; Indels 3; Gaps 3;

OY 5 CGNCGPSTPLPAMKGPREEIYVLPCTIRNTGTEADYATVVDPKSPQYCOVIHRLP 64
|||
Db 22 CKYGGPGAPPLAMSGPSEKLYVTAVTGTGIDKPDYATVVDSSPSVSIHRLP 81
|||||
OY 65 MPNLKDELHSHGWNCCSCFCGDSRTKSLVPLSLISSRIYVVDGSEPRAPKLHKYIEP 124
|||||
Db 82 MPYVGDDELHSHGWNCCSCFCGDSRTKSLVPLSLISSRIYVVDGSEPRAPKLHKYIEP 141
|||||
OY 125 KDIHAKELAFHLSHCLASGEVWISSLDGVKNGKGFVLLDGETFEVGTWERP 183
|||||
Db 142 KEIADKTGLAFPHAHCLATGEIILVSCIGDEGNAGKNGFLLDSD-FNKNRMEKPGHS 200
|||||
OY 184 APLOYDWMYQPRHNWMTSTEMAAPVLRDGENPADAVEAGLSHLYWDMQRIHEIVQ 243
|||||
Db 201 PLYGYDWMYQPRHNWMTSTEMAAPVLRDGENPADAVEAGLSHLYWDMQRIHEIVQ 260
|||||
OY 244 L-KDGLPILRFLNPDAAAGFCVCAALSTIQRFYKNEGTVSWVEKVIQVPPKKYGM 302
|||||
Db 261 LGPTGLPILRFLNPDAAAGFCVCAALSTIQRFYKNEGTVSWVEKVIQVPPKKYGM 320
|||||
OY 303 LPEMRELITDIILSLDDRFLYFSNMLHGLROYDISDPQRPRLTQGLFSGSYVKGCP 362
|||||
Db 321 LPEMRELITDIILSLDDRFLYFSNMLHGLROYDISDPQRPRLTQGLFSGSYVKGCP 380
|||||
OY 363 VLEDEELKQPEPLVYVKGKRVAGPOMIQLSDGKRLTYTTSLSYAMDQKQFPDLIRE 422
|||||
Db 381 AVGEDGNTFOPEVQIKGSLRGAPOMIQLSDGKRLTYTTSLSYAMDQKQFPDLIRE 440
|||||
OY 423 VMLQVDVDTVVGKGLKLPNCLVDFGKEPLGALAHRLRPGDCCSDIWI 472
|||||
Db 441 HIIDVDTEKGGTLINPDEFVDFGDEPDGSLHMKRYPGDCSTDIWI 490
|||||

RESULT 4
E71401
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 11-Jan-2002
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Entlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomen
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneve, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chaitatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113

```



```

OY 178 EPPGGAAPLGYFWYQPRHNWISTEWAAPNVLROGFNPADYEAQYSGSHLYWDMQRHE 237
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 176 PADEKTVFENYDFWQPRHNWISTEWAAPNVLROGFNPADYEAQYSGSHLYWDMQRHE 235
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 238 IYQTLSDK--GLIPEIRFLNHPAAGFVGCALSTIQRY--KNEGSTSVKRYIOY 293
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 236 YQITDLPQPLALPLEVAFLEPTSEHAFVGCALSGIFRHPVEENSTHAATLVAET 295
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 294 PPKVYKWLPPMPGLITDILSLDRLFFSNWLGDLROYDISDPORRLTGLQFLG 353
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 296 PPKVYKWLPPMPGLITDILSLDRLFFSNWLGDLROYDISDPORRLTGLQFLG 355
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 354 STYKGGPVYLEDDELKSPPELVYKGVAGGPGQMIQSLDGKRLYTTSTYSAMKOF 413
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 356 SVHTESNRYVLEGEK--PIEALYKGRKIEGGPQMLQISLQKRLYTTSTYSAMKOF 412
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 414 YPDLREGSVMLQVVDYTKGGLKLPNPLVDFGKPEPLPALAHELRVYGGCSDSIWI 472
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 413 YPEHVKSGATMVQVNIIDPSGKMEINRDLIDFGKIEGSPYLAHEKRIPGGCTSDIWI 471
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 7

A99464

selenium-binding protein [imported] - Sulfolobus solfataricus

C:Species: Sulfolobus solfataricus

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001

C:Accession: A99464

R:Singh, R.K.; Confalonieri, F.; Zivanovic, P.; Allard, G.; Awayez, M.J.; Chan

Jong, J.; Jeffries, A.C.; Kozera, C.J.; Medina, Y.; Peng, X.; Thi-Ngoc, H.P.; Redder, F

aretz, R.A.; Ragan, M.A.; Sengen, C.W.; Van der Oost, J.

Submitted to GenBank, April 2001

A:Description: Sulfolobus solfataricus complete genome.

A:Reference number: A99139

A:Accession: A99464

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-466 <KUR>

A:Cross-references: GB:AE006641; NID:g13816219; PIDN:AAK42968.1; GSPDB:GN00155

C:Genetics:

A:Gene: SSO2860

C:Superfamily: Caenorhabditis elegans hypothetical protein Y37A1B.5

Query Match 35.4%; Score 901.5; DB 2; Length 466;

Best Local Similarity 40.5%; Pred. No. 2.6e-66;

Matches 192; Conservative 83; Mismatches 154; Indels 45; Gaps 12;

```

OY 12 YSTPLEAMGPREEIYVLPICIRNMGTEADYATVDVDPKSPQYCOVIRHLPMPNLKOE 71
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 17 YSPKMAKSPEDLAVACLTGTGINRPDIAYVDVDPKSEETYSKYTHKVELSYVDE 76
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 72 LHHSGMNTSCSF---GSTKSRKLVLPISLISRIYVVDVSGSEPPAPKLHVEIPKDIH 128
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 77 LHHFGMNCSSALCPNGRNFERRFLVVGRLSSRIYIDTKLNRPQPIVATIEPEVK 136
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 123 AKCELAFLHTSHCLASGEVMISSLDVKGNGKGFVLLDGEFFEVKGTWERGGAPLGY 188
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 137 KVTGYSRHTLVACGPDG--IYISAFGENEGEGGILMDHVSFEPLGKWEIDRSQYLA 195
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 189 DEWYQPRHNWISTEWAAPNVLROGFNPADYEAQYSGSHLYWDMQRHEIYQTLSDK 247
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 196 DEWYQPRHNWISTEWAAPNVLROGFNPADYEAQYSGSHLYWDMQRHEIYQTLSDK 254
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 248 LIPEIRFLNHPAAGFVGCALSTIQRY--KNEGSTSVKRYIOY 293
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 255 RALPEIRFLNHPAAGFVGCALSTIQRY--KNEGSTSVKRYIOY 293
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 303 LPE-----MGULTDILSLDRLFFSNWLGDLROYDISDPORRLTGLQFLGSI 355
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 313 LPEILKPKRAVPLVTDILSLDRLFFSNWLGDLROYDISDPORRLTGLQFLGSI 372
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 356 VKGGPVYLEDDELKSPPELVYKGVAGGPGQMIQSLDGKRLYTTSTYSAMKOF 413
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 373 HRA-----DHP-----SDPKLTGAPQMIETISROGKRYVYVNSLISTWDMQYFP 415
OY 416 DLIREGSVMLQVVDYTKGGLKLPNPLVDFGKPEPLPALAHELRVYGGCSDSIWI 472
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 416 EBLKMMWKLKLNPNP---GGLDVDEKEFFVDEGE-----ARSHQVRLRGDASSD 461
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 8

T24193

hypothetical protein R11G10.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T24193

R:Burton, J.

Submitted to the EMBL Data Library, July 1996

A:Reference number: Z19851

A:Accession: T24193

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-576 <WIL>

A:Cross-references: EMBL:Z19851; PIDN:CA01239.1; GSPDB:GN00023; CESP:R11G10.2

A:Experimental source: clone R11G10

C:Genetics:

A:Gene: CESP:R11G10.2

A:Map position: 5

A:Introns: 41/3; 183/3; 225/1; 284/1; 364/3; 415/3; 491/2; 517/3

Query Match 26.9%; Score 685.5; DB 2; Length 576;

Best Local Similarity 35.6%; Pred. No. 2.3e-48;

Matches 165; Conservative 87; Mismatches 155; Indels 53; Gaps 16;

```

OY 24 EIVYLPICIRNMGTEADYATVDVDPKSPQYCOVIRHLPMPNLKDELHSGMNTSCSF 83
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 89 DELFAIVCCPHSIGYER-DKIALVDLPSTSEFTYILSEVHLTSGNDEPGRMMNAKSAS 147
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 84 FQDSKTK-SRTKLVLPISLISRIYVD-----VSEPPAPKLHVEIPKDIHAKCELA 135
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 148 LGEMNKFEVRNRIYPCMNMSGKITYIAFENEKLMIEKEIRNDELIR---KDV--SCPYAV 201
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 136 LHTSHCLASGEVMISSLDVKGNGKGFVLLDGEFFEVK---GTWERGGAPLGYDFW 191
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 202 --RSLPKAGAPVHVSITLDRFGNGKDFILLDRIVEVRAKSEPFSDYCG-----DFS 253
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 192 YQPRHNWISTEWAAPNVLROGFNPADYEAQYSGSHLYWDMQRHEIYQTLSDK--DG 247
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 254 LQPRHNWISTEWAAPNVLROGFNPADYEAQYSGSHLYWDMQRHEIYQTLSDK--DG 313
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 248 LIPEIRFLNHPAAGFVGCALSTIQRY--KNEGSTSVKRYIOY 293
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 314 SLVIVCFELHNADCNHAFISAIGSSIFHLHNNITLTKMAADRAVAHPLKVENMQSD 373
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 307 PGLITDILSLDRLFFSNWLGDLROYDISDPORRLTGLQFLGSIYVGGPVQV--- 363
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 374 PALLTDMITSMDDRLVYCGFLHGLVLMFRDIODPRVSLHGLKINGGIFDSPEVRITS 433
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 364 --LEDE-----ELKSQPELVYKGVAGGPGQMIQSLDGKRLYTTSTYSAMKOF 416
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 434 NAMEBRMWLPETRSPL-----RGTKFRGPGALMQLSDGKGLYVCSFYKAMDQOFP 488
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 417 LIREGSVMLQVVDYTKGGLKLPNPLVDFGKPEPLPALAHELR 460
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 489 LISDGGQMIYVDI--VDDEMOLNEKEFLIDMK---PSLIONYR 526
      1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 9

T32635

phosphoprotein phosphatase (EC 3.1.3.16) F4268.8 [similarity] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 02-Nov-2001

C:Accession: T32635

R:Galling, S.; Holmes, A.

Submitted to the EMBL Data Library, December 1997

A:Description The sequence of C. elegans cosmid F42G8.
A:Reference number: Z21203
A:Accession: F32635
A:Status: translated from GE/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-487 <GAT>
A:Cross-references: EMBL:AF038618; PIDN:AAB92072.1; GSPDB:GN00022; CESP:F42G8.8
A:Experimental source: strain Bristol N2; clone F42G8
A:Comment: This sequence shares a domain (approximately residues 378-487) with mammalian
C:Genetics:
A:Gene: CESP:F42G8.8
A:Map position: 4
A:Introns: 27/3; 97/2; 122/3; 203/3; 236/2; 274/3; 302/3; 377/1
C:Superfamily: Caenorhabditis elegans phosphoprotein phosphatase F42G8.8; phosphoesterase
C:Keywords: iron; metalloprotein; phosphoric monoester hydrolase; serine/threonine-specific
E:30-314/Domain: phosphoprotein phosphatase homology <PP>
F:58-137/Domain: phosphoesterase core homology <PEC>
F:64/66; 103/Binding site: iron (Asp, His, Asp) #status predicted
F:103;135;188;273/Binding site: zinc (Asp, Asn, His, His) #status predicted
F:106;116;297/Active site: Asp, His, Tyr #status predicted
F:107;246/Binding site: substrate phosphate (Arg) #status predicted

Query Match	14.4%	Score 366;	DB 1;	length 487;
Best Local Similarity	61.7%	Pred No. 4e-22;		
Matches 71;	Conservative 18;	Mismatches 22;	Indels 4;	Gaps 2;

[illegible]

RESULT 10
567660
hypothetical protein YDL117w - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: hypothetical protein D2275
C:Species: *Saccharomyces cerevisiae*
C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 29-Oct-1999
C:Accession: S67660
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.; Wagner, G.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67655
A:Accession: S67660
A:Molecule type: DNA
A:Residues: 1-885 <RIF>
A:Cross-references: EMBL:Z74165; NID:q1431171; PID:e253229; PID:q1431172; GSPDB:GN00004;
A:Experimental source: strain S288C
C:Genetics:
A:Gene: MIPS:YDL117w
A:Map position: 4L

Query Match	4.48;	Score 111;	DB 2;	Length 885;
Best Local Similarity	20.58;	Pred. No. 1;		
Matches 93;	Conservative 57;	Mismatches 153;	Indels 150;	Gaps 24

```

0Y      31  C1YRNTGTAEADYD1ATFADVPBPKSPQYQCVIHLRPMPLKDELHHS -GMWTKSSC----- 83
Db      264  CSSEHSFPLMDFEATSAGSPARKKYAO-----SFSDSLORSONAMCSTKINDSOE 315
0Y      84  FGDSTKSR-----TKLPLSPLSSRIYVYVDSGEPAPKILHVI----- 122
Db      316  FGDSMASSRNKMODILRKILIPRR-NTNIYSSV-SPPKSPAYRKLDPIDQNLNSTATP 373
0Y      123  -EPKD-IHAKCEL-----AFLTSHCLASEVMYISSLDYK 156
Db      374  DEARDWIWVKCHLNRARTLTGDKHPRYMRALLENRDLLIHPQDSLYNG---LNNNEVK 429
0Y      157  GNGKGFEVLLDGEFFEV-----KGWERBGAAGAPLGYDFW---YQPRHNWIS----- 201

```

```

Db      430 GNTKFG--LVDELAELEINIEYIDKMTWKRCIRDGTMTLLDSMAQTTFSARRSTYLEKRI 487
OY      202 -----TENA-----PNVLRD-GENPRADVAGLGYSN--YVMDQMHLEIVQLSLKDGILP 251
Db      488 YIECTEMFALTDNDGTSFSAEPONLEKILTRKCTRPYELTKRKLIANSIGTCEIV-- 545
OY      252 EIRFLHNPDAAO-----GFVSCAL--SSTIÖREYKKNEGTSVSEKY 290
Db      546 -IGFLKTPSAINWEKKYNHCKMLRLVYNKEWRFIVILGNVTNPRIHEVNNR---KAKKA 600
OY      291 -----LOVPRKVKYKGMILPEMGLITLIDLLSDDRFLYFSNMILHDLROY 336
Db      601 ENSYFLMAPLEMYTHIRPREFEQIIVPSIDQLSATLTP-----LVPSPFKNELKLYK 654
OY      337 ISDPQRRLTGQLPGSIVKGGROYLDEBL 369
Db      655 FS-----TALSFLDESEITECSEIENDVEV 680

```

RESULT 11
B82303
conserved hypothetical protein VC0603 [imported] - *Vibrio cholerae* (strain N16961 ser C);Species: *Vibrio cholerae*
C;Date: 18-Aug-2000 #sequence,revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: B82303
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R. ;
charidson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers
1. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
A;Reference number: A82035; MUID:20406833
A;Accession: B82303
A;Status: Preliminary
A;Molecule type: DNA
A;Residues: 1-764 <HE>
A;Cross-references: GB:AE004145; GB:AE003852; NID:9655032; PIDN:AAF93770.1; GSPDB:GN
A;Experimental source: serogroup O1; strain N16561; biotype El Tor
C;Genetics:
A;Gene: VC0603
A;Map position: 1

Query Match	4.3%;	Score 109.5;	DB 2;	Length 764;
Best Local Similarity	21.4%;	Pred. No. 1.1;		
Matches 55; Conservative	33;	Mismatches 84;	Indels 85;	Gaps 11;

[illegible]

RESULT 12
F71253
hypothetical protein TM1450 - Thermotoga maritima (strain MSB)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:47:16 ; Search time 13.43 Seconds

(Without alignments)
1360.806 Million cell updates/sec

Title: US-09-841-758-1
Perfect score: 2550
Sequence: 1 MATKCGNCGPSTPLEAMK.....PALAHRLRYPGDCSDIWI 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query	Length	ID	Description
1	2435	95.5	472	1 SBP1_HUMAN	Q13228 homo sapien
2	2283	89.5	472	1 SBP2_MOUSE	Q63836 mus musculu
3	2254	88.4	472	1 SBP1_MOUSE	P17563 mus musculu
4	1571.5	61.6	490	1 SBP_ARATH	O23264 arabidopsis
5	685.5	26.9	576	1 SBP_CAEEL	Q21950 caenorhabdi
6	103.5	4.1	407	1 SYX_THIRE	P41256 thiodacillu
7	102.5	4.0	781	1 COAT_PAVHB	P07299 human parvo
8	99	3.9	845	1 BGLS_KLUMA	P07337 kluyveromyc
9	98	3.8	365	1 CXAR_HUMAN	P08310 homo sapien
10	97.5	3.8	566	1 PRIM_BPT3	P20315 bacterioph
11	97.5	3.8	1159	1 DP3A_VIRCH	P52022 vibrio chol
12	97	3.8	757	1 PRGR_BOVIN	P81265 bos taurus
13	95	3.7	1783	1 Y468_MYCE	Q49460 mycoplasma
14	95	3.7	3066	1 POLG_SBMVG	Q90069 s genome po
15	94	3.7	437	1 HGD_CAEEL	Q9Y041 caenorhabdi
16	94	3.7	514	1 MP11_MOUSE	P48964 mus musculu
17	93.5	3.7	369	1 DHTS_HUMAN	P49366 homo sapien
18	93.5	3.7	529	1 KRT1_HUMAN	O00522 homo sapien
19	93.5	3.7	619	1 Y817_ARCFU	O29441 archaeoglob
20	92.5	3.6	538	1 REL_MELGA	P01125 melalegria
21	92.5	3.6	598	1 REL_CHICK	P16236 gallus gall
22	90	3.5	536	1 YAGH_ECOLI	P77713 escherichia
23	90	3.5	597	1 DY13_ANTCR	O16960 anthocidari
24	90	3.5	692	1 SPE1_SOYBN	O39827 glycine max
25	89.5	3.5	2813	1 VWF_HUMAN	P04275 homo sapien
26	89.5	3.5	3391	1 POLG_DEN26	P29990 d genome po
27	89	3.5	731	1 CAT1_HALMA	O59651 halocaula
28	88.5	3.5	577	1 RECA_ECOLI	P18933 escherichia
29	88.5	3.5	790	1 RECA_MYCTU	P66345 mycobacteri
30	88.5	3.5	926	1 UURA_AQUAE	O66911 aquitex aeo
31	88	3.5	365	1 CXAR_MOUSE	P97732 mus musculu
32	88	3.5	1070	1 PTK7_HUMAN	Q13308 homo sapien
33	88	3.5	1365	1 GTFS_STRDO	P29336 streptococc

34	87	3.4	504	1 CPDA_MOUSE	P24456 mus musculu
35	87	3.4	575	1 RECT_ERWCH	P39603 erwilia chr
36	87	3.4	715	1 SC14_HUMAN	O92503 homo sapien
37	87	3.4	1318	1 VP14_EBV	P03179 Epstein-Bar
38	86.5	3.4	502	1 KPYK_LACIA	O07637 lactococcus
39	86.5	3.4	680	1 GAOA_DADCE	Q01745 dactylium d
40	86.5	3.4	822	1 PBPA_PSEAE	O07806 pseudomonas
41	86.5	3.4	854	1 LDLR_CRIGR	P35950 cricetus
42	86.5	3.4	2365	1 TOXB_CLODI	P18177 clostridium
43	86	3.4	455	1 UHPT_CHLPN	Q927n9 chlamydia p
44	86	3.4	457	1 FKBA_MOUSE	P30416 mus musculu
45	86	3.4	1752	1 RPBI_SCHPO	P36594 schistosach

ALIGNMENTS

RESULT ID	1 SBP1_HUMAN	STANDARD	PRT	472 AA.
AC	Q13228			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Selenium-binding protein 1.			
GN	SELENBP1 OR SBP			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Heart;			
RX	MEDLINE=97179296; PubMed=9027582;			
RA	Chang P.W.G., Tsui S.K.W., Liew C., Lee C., Wayne M.M.Y., Fung K.;			
RT	"Isolation, characterization, and chromosomal mapping of a novel cDNA			
RT	clone encoding human selenium binding protein."			
RL	J. Cell. Biochem. 64:217-224(1997).			
CC	-1- FUNCTION: NOT KNOWN: BIND SELENUM.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE SELENUM-BINDING PROTEIN FAMILY.			
CC	-1- This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL: U29091; AAB02395.1; -.			
DR	MIY: 604188; -.			
KW	Selenium.			
SQ	SEQUENCE 472 AA; 52313 MW; F484CF1CD68FC3B CRC64;			
Query Match	95.5%;	Score 2435;	DB 1;	Length 472;
Best Local Similarity	96.0%;	Pred. No. 6.5e-196;		
Matches 453;	Conservative 3;	Mismatches 16;	Indels 0;	Gaps 0;
QY	1 MATKCGNCGPSTPLEAMKGRREIYVLCIYRNTGTADYLATVVDPRKSPQCOVI 60			
DB	1 MATKCGNCGPSTPLEAMKGRREIYVLCIYRNTGTADYLATVVDPRKSPQCOVI 60			
QY	61 HRLPMNPKDELHSHGWTGSCFGDSTKSTKLVLPGLISSRIYVVDGSEPRAPKHK 120			
DB	61 HRLPMNPKDELHSHGWTGSCFGDSTKSTKLVLPGLISSRIYVVDGSEPRAPKHK 120			
QY	121 VIEPKDIAKCELAFLHSHSLAGSEWISSLGNGKGGKGFVLLDGEFEVGTWRP 180			
DB	121 VIEPKDIAKCELAFLHSHSLAGSEWISSLGNGKGGKGFVLLDGEFEVGTWRP 180			
QY	181 GGAAPLGYDFVQRYRHNMISTEWAARVNLRDGNPADVAGLGSHLYVDMORHETVQ 240			

```

Db 181 GGAAPGIDFWYQPRHNWVISTEMAAPNVLKRGFNADVEAGLYGSHLYVWMOQREIYQ 240
QY 241 TSLKDGILPLEIRFLHNDAAGFYGALSSTIQRFYKNEGTVSEKVIQVPPKKYK 300
Db 241 TSLKDGILPLEIRFLHNDAAGFYGALSSTIQRFYKNEGTVSEKVIQVPPKKYK 300
QY 301 WLLPMPGLITDILSLDRFLYFSNWLHGLRQYDISDPQRPRLTGOLFGLGSIVKGP 360
Db 301 WLLPMPGLITDILSLDRFLYFSNWLHGLRQYDISDPQRPRLTGOLFGLGSIVKGP 360
QY 361 VOYLEDDELKSGPEPLVYKGRVAGGPOMIQLSDGKRLYTTSYLSAMKQFYPDLIRE 420
Db 361 VOYLEDDELKSGPEPLVYKGRVAGGPOMIQLSDGKRLYTTSYLSAMKQFYPDLIRE 420
QY 421 GSVMLQVDVDTYKGLKLPNLFVDFGKPEPLPALAHLRYGGDCSSDIWI 472
Db 421 GSVMLQVDVDTYKGLKLPNLFVDFGKPEPLPALAHLRYGGDCSSDIWI 472

```

RESULT 2

SBP2_MOUSE SBP2_MOUSE STANDARD: PRT: 472 AA.

```

AC 063836;
DT 15-JUL-1999 (rel. 38, Created)
DT 15-JUL-1999 (rel. 38, Last sequence update)
DT 16-OCT-2001 (rel. 40, Last annotation update)
DE Selenium-binding protein 2 (56 kDa acetaminophen-binding protein)
DE (A056)
GN SELENBP2 OR LPSB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C; TISSUE=Liver;
RX MEDLINE=93201669; PubMed=8453708;
RA Lantear J., Fleming J., Walker M., Harrison P.;
RT "Different patterns of regulation of the genes encoding the closely
RT related 56 kDa selenum- and acetaminophen-binding proteins in normal
RT tissues and during carcinogenesis.";
RL Carcinogenesis 14:335-340(1993).
RN [2]
RP SEQUENCE OF 175-189; 196-220; 228-242; 290-296; 334-343 AND 399-408.
RX MEDLINE=92171951; PubMed=1540179;
RA Pumphord N.R., Martin B.M., Hinson J.A.;
RT "A metabolite of acetaminophen covalently binds to the 56 kDa
RT selenium binding protein.";
RL Biochem. Biophys. Res. Commun. 182:1348-1355(1992).
CC -1- FUNCTION: NOT KNOWN. BIND SELENIUM AND ACETAMINOPHEN.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: MAINLY EXPRESSED IN LIVER.
CC -1- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: S56599; AAB25841.2;
CC DR GMD: MGI:104859; Selenbp2.
CC KW Selenium.
CC SEQUENCE 472 AA; 52628 MW; C32FE819C4AD07CA CRC64;

```

Query Match 89.5%; Score 2283; DB 1; Length 472;
 Best Local Similarity 87.5%; Pred. No. 3; 3e-183;
 Matches 413; Conservative 28; Mismatches 31; Indels 0; Gaps 0;

```

QY 1 MATKCGNGCPGISTPLEAMKGPREEIYVPCITRYNTGTADPYLATVVDKSPQCYI 60
Db 1 MATKCTKCGPYPTPLEAMKGPREEIYVPCITRYNTGTADPYLATVVDKSPQCYI 60
QY 61 HRLPMPNKLDELHSGWNTCCSCFGDSTKSRKLLPGLMSSRIYVVDGSEPRAPKLK 120
Db 61 HRLPMPNKLDELHSGWNTCCSCFGDSTKSRKLLPGLMSSRIYVVDGSEPRAPKLK 120
QY 121 VIEPKDIHAKELAEFLHSHCLASGEWMTSLGDYKNGKGGFVLLDGETFEVKGWERP 180
Db 121 VIEASEIQAKCVNMTHTSHCLASGEVWVNTLGLDQNGKGSFVLLDGETFEVKGWERP 180
QY 181 GGAAPGIDFWYQPRHNWVISTEMAAPNVLKRGFNADVEAGLYGSHLYVWMOQREIYQ 240
Db 181 GGAAPGIDFWYQPRHNWVISTEMAAPNVLKRGFNADVEAGLYGSHLYVWMOQREIYQ 240
QY 241 TSLKDGILPLEIRFLHNDAAGFYGALSSTIQRFYKNEGTVSEKVIQVPPKKYK 300
Db 241 TSLKDGILPLEIRFLHNDAAGFYGALSSTIQRFYKNEGTVSEKVIQVPPKKYK 300
QY 301 WLLPMPGLITDILSLDRFLYFSNWLHGLRQYDISDPQRPRLTGOLFGLGSIVKGP 360
Db 301 WLLPMPGLITDILSLDRFLYFSNWLHGLRQYDISDPQRPRLTGOLFGLGSIVKGP 360
QY 361 VOYLEDDELKSGPEPLVYKGRVAGGPOMIQLSDGKRLYTTSYLSAMKQFYPDLIRE 420
Db 361 VOYLEDDELKSGPEPLVYKGRVAGGPOMIQLSDGKRLYTTSYLSAMKQFYPDLIRE 420
QY 421 GSVMLQVDVDTYKGLKLPNLFVDFGKPEPLPALAHLRYGGDCSSDIWI 472
Db 421 GSVMLQVDVDTYKGLKLPNLFVDFGKPEPLPALAHLRYGGDCSSDIWI 472

```

RESULT 3

SBP1_MOUSE SBP1_MOUSE STANDARD: PRT: 472 AA.

```

AC P17563;
DT 01-AUG-1990 (rel. 15, Created)
DT 01-AUG-1990 (rel. 15, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE Selenium-binding protein 1 (56 kDa selenum-binding protein) (SP56).
GN SELENBP1 OR LPSB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91029655; PubMed=2225343;
RA Bansel M.P., Mukhopadhyay T., Scott J., Cook R.G., Mukhopadhyay R.,
RA Medina D.;
RT "DNA sequencing of a mouse liver protein that binds selenum:
RT implications for selenum's mechanism of action in cancer
RT prevention.";
RL Carcinogenesis 11:2071-2073(1990).
CC -1- FUNCTION: NOT KNOWN. BIND SELENIUM.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER, KIDNEY AND, TO A
CC LESSER EXTENT, LUNG.
CC -1- SIMILARITY: BELONGS TO THE SELENIUM-BINDING PROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC DR EMBL: M32032; AAA40104.1;
CC DR PIR: S27878; S27878.
CC SWISS-2DPAGE; P17563; MOUSE.

```


DR Pfam; PF00579; tRNA-synt_1b; 1.
 DR PRINTS; PRO1040; TRNASYNTHYR.
 DR SMART; SM00363; S4; 1.
 DR PROSITE; PS00178; AA_PRNA_LIGASE_I; FALSE NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
 FT SITE 47 56 "HIGH" REGION.
 FT SITE 231 235 "KMSKS" REGION.
 FT BINDING 234 234 ATP (BY SIMILARITY).
 SO SEQUENCE 407 AA; 45860 MW; 7B73FBA252A2C6B CRC64;

Query Match 4.18; Score 103.5; DB 1; Length 407;
 Best Local Similarity 19.38; Pred. No. 0.41;
 Matches 85; Conservative 61; Mismatches 110; Indels 185; Gaps 20;

QY 107 VDVGSEPRAPK-----LHKVIEPKDIAHAKCELAFLHTSHCLASGEVMISSLDGKGN 158
 Db 41 IKLMDPAPLHIGAYVLLHKAHOFDL-----GRLFLVIGD----- 79
 QY 159 GKGGFVLIDG-----TFEVKGTWERPGGAAPLGYDFWYOP-----RHNVMISTE 203
 Db 80 -----FTAMIGDPTGKSVTRKALSREEVANAT-----YRPQYFKILDPERTEVMRSE 129
 QY 204 W-----AADNVLIDGFN-----PADV 219
 Db 130 WLGAALPEELIOIACYAVARMLERDDFNKRYSANQPIAHEFLYPLDGYDVAIKADY 189
 QY 220 EAG-----LYGSHLYVMDQKHIEYQULSLDKGLPLEIRFLHPDAAQGFVGCALS 271
 Db 190 ELGGTDORFNLVGRELOREXGOKPOLVIMPLEGL----- 226
 QY 272 STORFKNNEGTSVEKIVQVPPKRYGMLPEMPGLITITLSDRFLYFSNWLH-- 329
 Db 227 DGVRKMSKSLGNFLAVED-----PRAEMFGKIMS-----ISFLM-----WRYA 266
 QY 330 -----GDLROYDISDPRPR-----LTGOL-----FLGSIYKGGPVQVLEDEELKS 371
 Db 267 LLSHRAVEQTRLOKEAASGARNPREDIKLAGELVRFHGTAAQEHMIFLARFQHE 326
 QY 372 QPEPLVYKGRVAGGPPQMIQSLDGRKLYITTSLYXSANDKQFYDDLREGSVMLQ-----V 427
 Db 327 TPEELPQATKLSAEPRLSOLLV-----QVHLASTSEAMRK-----MKEGAVRVDMWRKV 377
 QY 428 DVDVFKGGLKLNPNFLVDGFK 448
 Db 378 DPATF-----LALDAVYLLDQFK 395

RESULT 7
 COAT_PAVHB STANDARD; PRT; 781 AA.
 ID COAT_PAVHB
 AC P07299;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-APR-1990 (Rel. 14, Last annotation update)
 DE Probable coat protein VP1.
 OS Human parvovirus B19.
 OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Erythrovirus.
 OX NCBI_TaxID=10796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ISOLATE AU;
 RX MEDLINE=86200451; PubMed=3701931.
 RA Shade R.O., Blundell M.C., Cotmore S.F., Tattersall P., Astell C.R.;
 RT "Nucleotide sequence and genome organization of human parvovirus B19
 isolated from the serum of a child during aplastic crisis";
 RL J. Virol. 58:921-936(1986).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

CC entries requires a license agreement (See <http://www.isb.ch/announce/>
 or send an email to license@isb.ch).
 DR EMBL; M13178; AAA66867.1; -
 DR PIR; A24299; VCPV19.
 DR InterPro; IPR001403; Parvo_coat.
 DR Pfam; PF00740; Parvo_coat; 1.
 KW Coat protein; Glycoprotein.
 FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 781 AA; 86015 MW; 8C6254DBD0576B07 CRC64;

Query Match 4.08; Score 102.5; DB 1; Length 781;
 Best Local Similarity 20.58; Pred. No. 1.2;
 Matches 105; Conservative 57; Mismatches 160; Indels 191; Gaps 23;

QY 11 GYSTPLEAMKGPREEIYVLC-----IYKNTGTEADPYL-----ATVDPKSPQYCOV 59
 Db 309 GYSTPWRKLDNALNLFSPLEFQHLIENTYSLAPDALTYTISELAVKDYDTKGGVQV 368
 QY 60 IHRLPMPNLKDELHSGMNTSCSGFSDTSKRTKLVPLSLISSRI--YVVDGSEPRAP 116
 Db 369 T-----DSTGCR-----LCMLVDHEKKYVYVGGQDPLAP 399
 QY 117 KLHVIPEKDIHAKCELAFLHTSHCLASGEVMISSLDGKNGKGG-----FVL 165
 Db 400 EL-----PIWYFFPQYAYL-----TVGDNVNTQGISGSKKLASESESAFYV 440
 QY 166 LDGFEVFKGTWERPGGAAPLGYDFWYQPRHNVMISTEMAPNVLRDGFNPADVAGLYG 225
 Db 441 LEHSSFOILGT-----GGTASMSYKFPVPPENLE-----GCSQHFYEKNP-----LYG 485
 QY 226 SHLYVMDW-----QRHE-----IYQTLSLKDQ----- 247
 Db 486 SRLGVPDTLGSDPKFRSLTHEDHAIQPNEMFPGVLNVSYSRKEDSSNTGAKALTLGST 545
 QY 248 -----LIPLEI-RLHNPDAQGFVGCALSTIÖRFYKNGGTSVEKVIQVPPKK 297
 Db 546 GTSQNTRLSLRPGVSPQYHMDKVTYGINALSHGQTFYGNADKEDYEOGVGRFPNEK 605
 QY 298 VKGMLPEMPGLITDILLSDRFLYFSNMLHGLROYDISDPRRLTGOLF----- 350
 Db 606 EQ-----LKQDGL-----NMHTTFPN--KGTQQT-TQIREPLVGSVMRRALHY 649
 QY 351 -----LGGSIYKGGPVQVLEDEELKSQPEPLVYKGRVAGGPPOM 389
 Db 650 ESQLSKRIPNLDSFKIQFALGWMGLHQPPQIF-----LKILPQSGPIGSIKSMGITTL 705
 QY 390 IQSLSDGKRLYITTSL-----YSAMDKO--FYP 415
 Db 706 VQYAVGIMTYMTFKLGPRAKATGRMNQPPGYVP 738

RESULT 8
 BGLS_KLUMA STANDARD; PRT; 845 AA.
 ID BGLS_KLUMA
 AC P07337;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 01-APR-1993 (Rel. 25, Last annotation update)
 DE Beta-glucosidase precursor (EC 3.2.1.21) (Gentiobiase) (Celllobiase)
 DE (Beta-D-glucosidase glucosylhydrolase).
 OS Kluyveromyces marxianus (Yeast) (Kluyveromyces fragilis).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
 OX NCBI_TaxID=4911;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 12424;
 RX MEDLINE=88210533; PubMed=2835179;

RA Raynal A., Gerbaud C., Francinques M.C., Guerinneau M.;
 RT "Sequence and transcription of the beta-glucosidase gene of
 RL Kluyveromyces fragilis cloned in Saccharomyces cerevisiae.";
 CC Curr. Genet. 12:175-184(1987).
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of terminal, non-reducing beta-D-
 CC glucose residues with release of beta-D-glucose.
 CC -1- PATHWAY: CELLULOSE DEGRADATION.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X05918; CAA29353.1; -;
 DR PIR: A29148; GLVK.
 DR InterPro: IPR001764; Glyco_hydro_3.
 DR InterPro: IPR002772; Glyco_hydro_3C.
 DR Pfam: PF00933; Glyco_hydro_3; 1.
 DR Pfam: PF01915; Glyco_hydro_3_C; 1.
 DR PRINTS: PR00133; GLHYDRASE3.
 DR PROSITE: PS00775; GLYCOSYL_HYDROL_F3; 1.
 DR PROSITE: PS00775; Cellulose degradation; Glycoprotein; Signal.
 KM Hydrolyase; Glycosidase; Cellulose degradation; Glycoprotein; Signal.
 FT SIGNAL 1
 FT CHAIN ? 845
 FT ACT_SITE 225 225 BY SIMILARITY.
 FT CARBOHD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CAROHD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 845 AA; 93916 MW; 78D4C150D5992B5 CRC64;

Query Match 3.9%; Score 99; DB 1; Length 845;
 Best Local Similarity 19.8%; Pred. No. 2.6;
 Matches 96; Conservative 47; Mismatches 123; Indels 218; Gaps 24;

2 ARKCGNGGYSSTPLEAMKGPREEIYLPCIRNNGTEADVLAVD----- 48
 DB 408 AKNAGLIAKFFYNPVE-----ERSEDEPPEPVTKVNSVHLFDFKHEK 451
 QY 49 VDPKSPQYCOVTHRLPMN-----LKDEL-----HHSWNCTSCFCG 85
 DB 453 VDPKPNPFVTLTGQYVQDEGDYIFSLQVYSGLEFLINDELIDQKHNQ--ERGSFCFG 509
 QY 86 DSTKSRKIVLPSLISSRIYVDV--GSEP-----RAPKLKVIIEPKDIHAK 130
 DB 510 AGTKERTKTL--TLKKGVYNYRVVEGSGPTSGLVGEFGAGCGAGVIAIKADDEIRNA 567
 QY 131 CELATLHSHCLASGEVMISLGDVKGNGKGFVLLDGTFFVKGWTERPG----- 181
 DB 568 AELAKHKAVLIIG-----LNEWEETEGYDRENMDLP 600
 QY 182 -----GAAPLGYDF-----WYQPRH----- 196
 DB 601 KRTNELVRAVLKANPNYIVNOSGTFPERPWEANALVQANGNELGNALADVLGYGV 660
 QY 197 --NVMISTEWAAPNVLRDGFNPA---DVEAG--LYGSHLYVMDORHEIVOTLSLKDGL 248
 DB 661 VPMGKLSLSM--PKLQD--NPAPLNFKTEFGRAVYVEDIEF--GYRYEKLQ----- 707
 QY 249 IPLEIRFLHNPDAAGFCVCCALST-----IOFFYKNEGTSVEKEYI 291
 DB 708 -----RKVAFP-----FGYGYSTTFEELIDSEFVYTDKIDISVDVKNKGTGDFAGSEVY 756
 QY 292 QV-----PPKKVGMLLPEM--PG--LITDILLSDRFLYFSNWL--HGDLRO 334
 DB 757 QVYFSALNSKVSRYKELKGFKEKVLHEPEKEKTVNIELEKDAISYFNEELCKMWAVEAGE 816

QY 335 YDIS 338
 DB 817 YLVS 820

RESULT 9
 CXAR_HUMAN STANDARD; PRT; 365 AA.
 ID CXAR_HUMAN
 AC P78310; 000694;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Coxsackievirus and adenovirus receptor precursor (Coxsackievirus B-
 DE adenovirus receptor) (hCAR) (CVB3 binding protein).
 GN CXADR OR CAR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCB1_Taxid=9606;
 OK NCB1_Taxid=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97190109; PubMed=9036860;
 RA Bergelson J.M., Cunningham J.A., Droguett G., Kurt-Jones E.,
 RA Krithivas A., Hong J.S., Horvitz M.S., Crowell R.L., Flahberg R.W.;
 RT "Isolation of a common receptor for Coxsackie B viruses and
 RT adenoviruses 2 and 5.";
 RL Science 275:1320-1323(1997).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97250541; PubMed=9096397;
 RA Tomko R.P., Xu R., Philipson L.;
 RT "hCAR and hCAR: the human and mouse cellular receptors for subgroup C
 RT adenoviruses and group B coxsackieviruses.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:3352-3356(1997).
 RN (3)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20008750; PubMed=10543405;
 RA Bowles K.R., Gibson J., Wu J., Shaffer L.G., Towbin J.A.,
 RA Bowles N.E.;
 RT "Genomic organization and chromosomal localization of the human
 RT Coxsackievirus B-adenovirus receptor gene.";
 RL Hum. Genet. 105:354-359(1999).
 RN (4)
 RP SEQUENCE FROM N.A.
 RA Anderson C.W., Kieleczawa J., Dunn J.J., Freimuth P.;
 RT "Sequence and expression of CXADR, the human gene for the
 RT coxsackievirus and adenovirus receptor.";
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: SERVES AS A RECEPTOR FOR GROUP B COXSACKIEVIRUSES AND
 CC SUBGROUP C OF ADENOVIRUSES (AD2 AND AD5).
 CC -1- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y07593; CAA68868.1; -;
 DR EMBL: U90716; AAC51234.1; -;
 DR EMBL: AF169366; AAF05908.1; -;
 DR EMBL: AF169360; AAF05908.1; JOINED.
 DR EMBL: AF169361; AAF05908.1; JOINED.
 DR EMBL: AF169362; AAF05908.1; JOINED.
 DR EMBL: AF169363; AAF05908.1; JOINED.
 DR EMBL: AF169364; AAF05908.1; JOINED.
 DR EMBL: AF169365; AAF05908.1; JOINED.
 DR EMBL: AF200465; AAF24344.1; -;
 DR MIM: 602621; -;

DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig_2.
 DR SMART: SM00410; Ig_Like; 1.
 DR SMART: SM00408; IgC2; 1.
 KW Immunoglobulin domain; Receptor; Transmembrane; Glycoprotein; Signal;
 Repeat.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 365 COXSACKIEVIRUS AND ADENOVIRUS RECEPTOR.
 FT DOMAIN 20 237 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 238 258 POTENTIAL.
 FT DOMAIN 259 365 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 34 127 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 155 219 IG-LIKE C2-TYPE DOMAIN 2.
 FT DISULFID 41 120 BY SIMILARITY.
 FT DISULFID 162 212 BY SIMILARITY.
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 365 AA; 40029 MW; AB01C6346CB7FE64 CRC64;

Query Match 3.88; Score 98; DB 1; Length 365;
 Best Local Similarity 17.4%; Pred. No. 1;
 Matches 82; Conservative 72; Mismatches 140; Indels 178; Gaps 21;

QY 13 STEPEAKKGPREEIVYPCIRYRNTGTEAPDYLAIVDYD-PSKPYQCVIHRHP- 65
 DB 23 TTPREMIKAKGEYAYLPCKF---TLSPEDQGPLDIEMLISPADNOKVDVVIILYSGDK 78
 QY 66 -----PNLDELHSGWNTCCSGCDSTKSRTKLVPSLISRIYVVDGS---EPRA 115
 DB 79 IYDYVDLKGRAVFTS-NDLKS--GDASINVTNLOLS-----DIGYQCKYKKA 125
 QY 116 P-----KLKVIYEPKDIHAKCELAFLHTSHCLASGEVWISSLGKNGKGGFVLLDGET 170
 DB 126 PGVANKKITHLVLPKPSGARC-----YVDGSEEI-----GSD 157
 QY 171 FEVATGWERPRGGAAPLYGDFWYOPRHNVMISTEWAAPNVLRDGFNPADVEAGLGSMLYV 230
 DB 158 FKIR--CEPKEGSLPLOYE-WQLSDSOKMPTSLA----- 190
 QY 231 WDWQRHEIVOT-LSLKDGLIPLEIRFLNPDAAQGFVGCALSTIIOEFYKNGGTWSEVK 289
 DB 191 -----EMTSVLSYKNA-----SSEYSGTYSCTYRNRY-----GSDQCLLR 226
 QY 290 VIQYPRKKVKGWLLPEMPGLITDLSL-----DGRFLYSNNMLHGLROYDI 337
 DB 227 LNVVPPSPNKAAGL---AGALIGTLALALIGLIIFCCRRKKRREKEVEVHNDIRE-DV 281
 QY 338 SDPQRPLTGLQFLGSGIVGSPV-----QVLEDELKSGQREPLVYKGRVA 384
 DB 282 PPRPSKSTARSYIGSNHSLGSMSPSNMEGYSTQYNQVSEDFERTPOSPTLPKAKVA 341
 QY 385 GGPQMIGLSIDGKRLYITTSLSYAMDKQFYFDLIREGVSVMLOVDVDTYVGGGL 436
 DB 342 A-----PNLSRMGAIPMPIDPAQKDGSI 364

RESULT 10
 PRIM_BPT3 STANDARD; PRT; 566 AA.
 AC P20315;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE DNA primase/helicase (EC 2.7.7.-).
 GN 4.
 OS Bacteriophage T3.
 OC Viruses: dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
 OC T7-like phages
 OC NCBI_Taxid=10759;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-LORIA;
 RX MEDLINE-90133923; PubMed-2614843;
 RA Beck P.J., Gonzalez S., Ward C.L., Molinex I.J.;
 RT "Sequence of bacteriophage T3 DNA from gene 2.5 through gene 9."
 RL J. Mol. Biol. 210:687-701(1989).
 CC -1- FUNCTION: DNA PRIMAASE SYNTHESIZES RNA PRIMERS NECESSARY FOR
 CC REPLICATION BY DNA POLYMERASE. IT ALSO FUNCTIONS AS AN ATP-
 CC DEPENDENT UNWINDING PROTEIN. HAS A 5'->3' HELICASE ACTIVITY;
 CC 4A ALSO HAS A DNA PRIMAASE ACTIVITY (BY SIMILARITY).
 CC -1- ALTERNATIVE PRODUCTS: CHAINS A (4A) AND B (4B) ARE THE PRODUCTS OF
 CC GENE 4 AND ARE PRODUCED FROM ALTERNATIVE INITIATION CODONS (BY
 CC SIMILARITY).
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X17255; CAA35135.1; -.
 DR EMBL: X17255; CAA35136.1; ALT_INIT.
 DR PIR: S07508; S07508.
 DR InterPro: IPR002936; Toprim.
 DR Pfam: PR01751; Toprim; 1.
 DR SMART: SM00493; TOPRIM; 1.
 KW DNA replication; DNA-directed RNA polymerase; Primosome; Late protein;
 KW Alternative initiation; Zinc-finger; Metal-binding; Helicase;
 KW ATP-binding; Transferase.
 FT CHAIN 1 566 DNA PRIMAASE/HELICASE, A CHAIN (BY
 FT SIMILARITY).
 FT CHAIN 63 566 DNA PRIMAASE/HELICASE, B CHAIN (BY
 FT INIT_MET 63 63 SIMILARITY).
 FT ZN_FING 15 37 FOR B CHAIN (BY SIMILARITY).
 FT C4-TYPE: MAY RECOGNIZE PRIMING SITES
 FT (POTENTIAL).
 FT NR_BIND 311 318 ATP (POTENTIAL).
 SQ SEQUENCE 566 AA; 62740 MW; 7553913043DDEBE CRC64;

Query Match 3.88; Score 97.5; DB 1; Length 566;
 Best Local Similarity 21.2%; Pred. No. 2;
 Matches 83; Conservative 47; Mismatches 135; Indels 127; Gaps 19;

QY 128 HAKCELAFLHTSHCLASGEVWISSLGKNGKGGFVLLDG-----ETFEV 173
 DB 12 HAPCE-----NC-----GSSDGSNSYSDGHEWCFCEHVPANEREA 49
 QY 174 K-GTWERPRGGAAPLYGDFW---YOPRHNVMISTEWAAPNVLRDGFNPADVEAGLY---- 224
 DB 50 KLSTRRTTGSKPSTYDVMWNGDSNGRISDLTANGISKETCOKAGWYLAKVDNRMYOYAD 109
 QY 225 -----GSH-----LYVMDQRHEIVQTLISLKDGLIPLEIR 254
 DB 110 YRDNGSIVSGKVRDKNFKTTGSHKSDALFLKHLMSGKR-KLYVGEIGDALTYMELQ 168
 QY 255 FLNPDAAQGFVGCALST-----IQRYFKNEGGSVEKYIYQV-PPKKVK 299
 DB 169 DCKYPVVSLHGSAARKTKCANAYEYDFOEQIILMEDMDAGRAVEEAQVLPACKVR 228
 QY 300 GWLLPEMPGLITDLSLDRFLTFSNW-----LHGDLRQYDSDQRPRLT-----GQ 346
 DB 229 VAVLPCKDA--NECHINGEDATILEQYWNANPWYRQDVSAISLKDVKKAMTSEDAVGL 286
 QY 349 LFLGGSIVKGGPVQVLEDELKSGQREPLVYKGRVAGGPOMIQLSLDGKRLYITTSLSYA 408
 DB 287 LFDS-----CGGLNDRILGARGGEVW---VTSGSGH-----GKSTFVHQAL-A 327
 QY 409 WDKQFYFDLIREGVSVMLOVDV-DTVKGGGLKN 439
 DB 328 WGRKMGK---RVGLAMLEESVEDFIQDMGIN 356

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: I04797; AAC41620.1; -
 DR EMBL: X81371; CA57136.1; -
 DR InterPro: IPR003599; Ig.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR003600; Ig_Like.
 DR Pfam: PF00047; Ig; 5.
 DR SMART: SM00409; Ig; 4.
 DR SMART: SM00410; Ig_Like; 1.
 KW Immunoglobulin domain; Repeat; Transmembrane; Glycoprotein; Signal;
 KW Polymorphism; Phosphorylation; Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 757
 FT DOMAIN 19 599
 FT TRANSMEM 633 653
 FT DOMAIN 654 757
 FT DOMAIN 19 132
 FT DOMAIN 133 238
 FT DOMAIN 239 350
 FT DOMAIN 351 462
 FT DOMAIN 463 575
 FT DISULFID 40 110
 FT DISULFID 56 64
 FT DISULFID 152 220
 FT DISULFID 257 324
 FT DISULFID 271 279
 FT DISULFID 370 440
 FT DISULFID 384 394
 FT DISULFID 481 543
 FT DISULFID 485 519
 FT DISULFID 495 519
 FT CARBOHYD 83 83
 FT CARBOHYD 420 420
 FT CARBOHYD 468 468
 FT MOD_RES 727 727
 FT VARSPPLIC 129 346
 FT VARIANT 29 29
 FT VARIANT 142 142
 FT VARIANT 404 404
 FT VARIANT 413 413
 FT VARIANT 435 435
 SO SEQUENCE 757 AA; 82434 MW; DCED67FDD6A6E6C6 CRC64;
 Query Match 3.8%; Score 97; DB 1; Length 757;
 Best Local Similarity 22.4%; Pred. No. 3.3;
 Matches 89; Conservative 46; Mismatches 131; Indels 132; Gaps 25;
 QY 4 KCG-----NCPRGSTPLEAKKPRE-----EIVYLPCITNTGTEADYATV 47
 DB 109 KCGIGISSRLNPDVSLLEVQDPAQASHAHVYVDLGRTYTINCFRANSEKRSKLC-- 166
 QY 48 DVDPKSPQVC-QVIHR--LMPNKLDELHHS--GMNCCSCFGDSTSRKIVLPSSLIS 102
 DB 167 ---KRTIQDCFQVYVDSIGVYSNKYKRAHISILGINT-----LVFSVYIN 208
 QY 103 RI-----YVVDVSGEPRAPKLH--KVLEPK-----DIHAK-----CEIA 134
 DB 209 RKLSDAGMYVCOAGDDAKADKINIDQVLEPEBELYGLLRSSVFDCSLGPEVANVPK 268
 QY 135 FLHTSHCLASGEVWISSLGIVKNGKGFVLL--DGTFEYKGTWEPG-----GAAP 185
 DB 269 FLOCKKNGACNVYINTLAKKADFOGRIVSVKDNQGVFVHTSLKEDDAGRIVCGAOP 328

QY 186 LG--YDFWQPRRHVMISTEWAP-----NVLRDG-----ENPADVAGLYGSHLYV 230
 DB 329 EGEGQDMPQAMOLPNEEELAIIPASPSYKGVGGSVYSCPNPDANSKATWCH--- 385
 QY 231 WDWGHEIVOTLSIKDLIPLE---IRFLHNP-----DAAGFVGALST 273
 DB 386 WEAQONCRCPRLVRSRLIKEQYEGRLALLTPENGTYVILNQLTQDQGFYWCV---- 441
 QY 274 IQRFYKNEGGT-W--SYE-KYIQ-----VPPKKYKGL 302
 DB 442 -----TDGTRWISTVELKVGGEPSLKVPKNTAWL 473
 RESULT 13
 Y468_MYCGE STANDARD; PRT; 1783 AA.
 AC Q49460;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUL-1999 (Rel. 38, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Hypothetical protein MG468.
 GN MG468.
 OS Mycoplasma genitalium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxId=2097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=96026346; Pubmed=7569993;
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP REVISIONS.
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fitchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,
 RA Nguyen D.T., Uterback T.R., Sauder D.M., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 879-985 FROM N.A.
 RC STRAIN-ATCC 33530 / G-37;
 RX MEDLINE=94075230; Pubmed=8253680;
 RA Peterson S.N., Hu P.-C., Bort K.F., Hutchison C.A. III;
 RA "A survey of the Mycoplasma genitalium genome by using random
 RT sequencing";
 RL J. Bacteriol. 175:7918-7930(1993).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
 CC -1- SIMILARITY: SOME, TO MG064.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U39728; AAC72488.1; -
 DR EMBL: U01808; AAD12339.1; -
 DR TIGR: MG468; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 16 36
 FT TRANSMEM 917 937
 FT TRANSMEM 967 987
 FT TRANSMEM POTENTIAL.
 FT TRANSMEM POTENTIAL.

FT TRANSMEM 1010 1030 POTENTIAL.
 FT TRANSMEM 1084 1104 POTENTIAL.
 FT TRANSMEM 1680 1680 POTENTIAL.
 FT TRANSMEM 1709 1729 POTENTIAL.
 FT TRANSMEM 1730 1750 POTENTIAL.
 FT TRANSMEM 1752 1772 POTENTIAL.
 SQ SEQUENCE 1783 AA; 200168 MW; 87BD575AEC2E374B CRC64;

Query Match 3.7%; Score 95; DB 1; Length 1783;
 Best Local Similarity 21.7%; Pred. No. 16;
 Matches 50; Conservative 33; Mismatches 85; Indels 62; Gaps 10;

QY 240 QTLSTKDLIPLEIRLHNPDAAGFVGCASTIORFYKNG-GTWSYEVYQVPPKKV 298
 DB 1185 QPKRADG-TELELGNLLP-----SYGGAQOQNTDENFRRHSLSKMIDPFIIRGSGSI 1239
 QY 299 KGWLLPE--MGLITDILSLDRLFRYSNMLHGLRLROYDISDPORPLTGLGSLIV 356
 DB 1240 NFWETVERKSIPKQITQLLSSASDQFL----- 1265
 QY 357 KGPQVQVEDELKSGPPELVYKGRVAGGPQMIQLSDGRKLYITTSLSYAMDKQFYPD 416
 DB 1266 -----ISVLTDEFFNN-----LNAAGFLIR-NPRTNOIQLDASRYVLTITIDYFNGGVKFNDS 1316
 QY 417 LIRESGVLYQ--DVTYKGGKLNPNFLVDGKEPLGALAHEDRYPGG 464
 DB 1317 FL--SFMLKVGDFELAKQDSKLN-----FGIVPVDPAIEFYTYVVG 1357

RESULT 14
 POLG_SBMVG STANDARD; PRT: 3066 AA.

AC 090069;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Genome polypeptide (Comtains: N-terminal protein (P1); Helper component polypeptide (EC 3.4.22.-) (HC-PRO); Protein P3; 6 kDa protein 1 (6K1); Cytoplasmic inclusion protein (CI); 6 kDa protein 2 (6K2); Genome-linked protein (VPG); Nuclear inclusion protein A (NI-A) (NIA) (EC 3.4.22.44) (49 kDa proteinase) (49 kDa-PRO); Nuclear inclusion protein B (NI-B) (NIB) (RNA-directed RNA polymerase) (EC 2.7.7.48); Coat protein (CP)).
 OS Soybean mosaic virus (strain G2) (SMV).
 OC Viruses: ssRNA positive-strand viruses, no DNA stage; Polyviridae; Polyvirus.
 NC NCBI_TaxID=103931;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92356085; PubMed=1645142;
 RA Jayaram C., Hill J.H., Miller W.A.;
 RT "Complete nucleotide sequences of two soybean mosaic virus strains differentiated by response of soybean containing the Rsv resistance gene";
 RL J. Gen. Virol. 73:2067-2077(1992).
 CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT MAY BE INVOLVED IN REPLICATION.
 CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: Hydrolyses glutamyl bonds, and actively is further restricted by preferences for the amino acids in p6 - p1' that vary with the species of polyvirus, e.g. Glu-Xaa-Xaa-Tyr-Xaa-Gln+Ser or Gly) for the enzyme from tobacco etch virus. The natural substrate is the viral polypeptide, but other proteins and oligopeptides containing the appropriate consensus sequence are also cleaved.
 CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT INDIVIDUAL PROTEINS.

CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC -1- SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC -1- SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: S42280; AAB22819.2; -
 DR MEROPS: C04.001; -
 DR MEROPS: C06.001; -
 DR Interpro: IPR001410; DEAD.
 DR Interpro: IPR001650; Helicase_C.
 DR Interpro: IPR001730; peptidase_C4.
 DR Interpro: IPR001456; peptidase_C6.
 DR Interpro: IPR002540; poly_p1.
 DR Interpro: IPR001592; poly_coat.
 DR Interpro: IPR001205; RNA_pol_P3D.
 DR Interpro: IPR001254; Trypsin.
 DR Pfam: PF00270; DEAD_1.
 DR Pfam: PF00271; helicase_C_1.
 DR Pfam: PF00863; peptidase_C4_1.
 DR Pfam: PF00851; peptidase_C6_1.
 DR Pfam: PF00767; poly_coat_1.
 DR Pfam: PF01577; poly_p1_1.
 DR Pfam: PF00680; RNA_dep_RNA_pol_1.
 DR PRINTS: PRO0966; NIA_POTYVIRASE.
 DR SMART: SM00487; DEXDC_1.
 DR SMART: SM00490; HELIC_C_1.
 DR Hydrolase: Transferase; Thiol protease; RNA-directed RNA polymerase;
 KW Coat protein; Polypeptide; Covalent protein-RNA linkage; Helicase;
 KW ATP-binding.
 FT CHAIN 1 ? N-TERMINAL PROTEIN.
 FT CHAIN ? ? HELPER COMPONENT PROTEINASE.
 FT CHAIN ? ? PROTEIN P3.
 FT CHAIN ? ? 6 KDA PROTEIN 1.
 FT CHAIN ? ? CYTOPLASMIC INCLUSION PROTEIN.
 FT CHAIN ? ? 6 KDA PROTEIN 2.
 FT CHAIN ? ? GENOME-LINKED PROTEIN.
 FT CHAIN ? ? NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN ? ? NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN ? ? COAT PROTEIN.
 FT NP_BIND 1249 1256 ATP (POTENTIAL).
 FT SEQUENCE 3066 AA; 349538 MW; 2188A79DBD155399 CRC64;

Query Match 3.7%; Score 95; DB 1; Length 3066;
 Best Local Similarity 19.6%; Pred. No. 34;
 Matches 98; Conservative 52; Mismatches 150; Indels 200; Gaps 29;

QY 32 IYRNTGTER-----PDLATVVDVDPKSPQ-YQVYHRLP-MPNLKDELHSGMWTGSSC 83
 DB 1969 VYHNGDLAYFGKNTSEALKVDLPHRPTLCONSAIAGPREDDELKRTG----- 2021
 QY 84 FQDSKSTKLVLPISLRIVVDVGSSEPRAPKLHVIPEPDIAK-----CE 132
 DB 2022 -----LPQVYKSS-----DY-----PRA-KERVEMESKSYVKGRLDSIGSTILCQ 2061
 QY 133 LAFHTSHCLASGEWISSLDGKNGKGGFVLDGETFE-----VKGTW----- 177
 DB 2062 LTNSSDGH-----KETMF-----GVGGSFILTNGHLFRRNGLTVK-TWNGSEFVTHN 2109
 QY 178 -----ERPGGAPLG-YDWTYQPRHN-----VMISTE----- 203
 DB 2110 TTOLKIHFIOGRDYLILMPKDFPFGKRNLFQPKREERYCMGTNFOEKSIRATVES 2169
 QY 204 -----MAANVLRDGF-----NPADVAGLGYSHLVWMDQREIYQTLSTL 244
 DB 2170 SKILPGRGKSFHWITTYDGGCGGLPLVSVNDGHI-VGIHG-----LTSNDSE 2216

QY 245 KDLGLPL-----EIRFLNPDAAQGVCAALSTIOREYKNGTWSYEKVIQVPPK--- 297
 Db 2217 KNFEVPLTDFGEKYLEMAD-----NLSDMKHMFWEPSKIMGSLNINLEDDPKKEFK 2268
 QY 298 -----VKG-----WLLPEMPGLITDILLSLDRFL-----YFSNWL 328
 Db 2269 ISKLVSDFGNTVYVGRKRWVLDAMEGNLA--ACGQDSALVTKHYVKGKCPFAQYL 2326
 QY 329 HGDLROYDISP-----ORPRLTGQLFGSIYKGPQVLEDELEKSOPEPLVYKGRV 383
 Db 2327 SVNOEAKSFPEPLMGAVOPSRFLNKDAFRGFFKYNKPV-VLNEVDFQS-----PERA 2377
 QY 384 AGCPOMIOLSLDKR-LYIT 402
 Db 2378 VAGVKMMMERDFKECYVT 2397
 RESULT 15
 HGD_CAEEL STANDARD; PRT; 437 AA.
 AC 09Y041: 062087: 09NJJP3:
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Homogentisate 1,2-dioxygenase (EC 1.13.11.5) (Homogentisicase)
 GN Homogentisate oxygenase (Homogentisic acid oxidase).
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Pelodierinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-BRISTOL N2;
 RC Schmidt S.R., Werner E., Mueller C.R., Kress W.;
 RT "Cloning and characterization of the homogentisate 1,2-dioxygenase
 RT gene in A. thaliana and C. elegans.";
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Schmidt S.R., Werner E., Mueller C.R., Kress W.;
 RT "Sequence homology of HGO genes in eukaryotic organisms.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Kersting J.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP REVISIONS.
 RA Durbin R.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: Homogentisate + O(2) = 4-maleylacetate.
 CC -|- COFACTOR: IRON.
 CC -|- PATHWAY: Catabolism of tyrosine; third step.
 CC -|- SIMILARITY: BELONGS TO THE HOMOGENTISATE DIOXYGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U95181; AAD00776.1; -
 DR EMBL: AF136150; AAF61419.1; -
 DR EMBL: Z93778; CAB07848.2; -
 DR EMBL: AL034364; CAB07848.2; JOINED.
 DR EMBL: AL034364; CAA22255.2; -
 DR EMBL: Z93778; CAA22255.2; JOINED.

DR HSSP: Q93099; IEYB.
 DR Wormpep: W06D4.1; CE23602.
 KW Oxidoreductase; Dioxygenase; Iron; Phenylalanine catabolism;
 KW Tyrosine catabolism.
 FT METAL 336 336 IRON (BY SIMILARITY).
 FT METAL 342 342 IRON (BY SIMILARITY).
 FT METAL 372 372 IRON (BY SIMILARITY).
 FT CONFLICT 156 156 D -> N (IN REF. 1).
 FT CONFLICT 176 176 L -> P (IN REF. 1).
 FT CONFLICT 193 193 V -> G (IN REF. 1).
 FT CONFLICT 242 242 V -> G (IN REF. 1).
 FT CONFLICT 265 265 D -> E (IN REF. 1).
 FT CONFLICT 433 433 F -> Y (IN REF. 1).
 SO SEQUENCE 437 AA; 49238 MM; C14E7077C7CE9703 CRC64;

Query Match 3.7%; Score 94; DB 1; Length 437;
 Best Local Similarity 20.4%; Pred. No. 2.8;
 Matches 103; Conservative 55; Mismatches 152; Indels 194; Gaps 28;

QY 24 EIIYLPCTYRNTGTEAPDYLATVVDPKSPQYC----- 57
 Db 5 DELKTLGFGNEHATSDPRVDALPVGONSPOKSHGILAOJSGTAFARSONQNSWL 64
 QY 58 -----QVTHRLPMPNLKDELHHSWG-MTCSSCFGDSKSR-TKLVLPSLISSRIYVDVG 110
 Db 65 YRIPSVYHR-PFEAMKENDQH-WTNNESSIPNPNGYRKNPPPLT-KEGYTFVD- 117
 QY 111 SEPAAPKAYIEPRDKHAKCELAFLTSHCLAGE-VMISSIGD----- 154
 Db 118 -----MLYTCGGGDVYSRTGLA-IHOFSCNASMEHTAMVNSDGLIIVPQGALEITTT 170
 QY 155 -----VKNGKGFVLLDGETFEVKGTEMRPGGAPLG--- 187
 Db 171 EFGRLVNPQELAVIPQGISVAVKGRSR-GYLL-----EYGTNHLQPLDLPICANG 223
 QY 188 ---YDF---WYQPRHNVMISTEMAPN-----VLRDGFNPADYAGLYGSHL-YV 230
 Db 224 LANPRDFEAPVAMPED-----LDVEFTIKYQGSWFGAKGHSFPDV-VGMHGNVYVK 277
 QY 231 WDMQRHEIVOTLSLKDGLIPLLEIRFLNPDAAQ-----FY-----GCALSTIOREYK 279
 Db 278 YDLKFMVINTVSR-DHODPSIFTVLTAPSVKHGTALADFEVIFPPRMGCACANTFRPPYH 336
 QY 280 NEGWTWSYEKVIQVPPKVKWMLPEMPGLITDILLSLDRFLYFSNWLHGDLYQYDSD 339
 Db 337 RN-----CMSEYGLITTGCTEAKEGGFKPGGSLHSMATPH----- 372
 QY 340 PORPRLTGQLFGSIYKGP-----VOVLEDELEKSOPEPLVYKGRVAGGPO--MIOLS 393
 Db 373 -----GDFMCFEMASNADLKPO-----RYABETGSMFMRESS 404
 QY 394 LDGKRLVYTT-SLYSANDKQFYPD 416
 Db 405 LN---MAITMNAVYQVNDKDYKD 425

Search completed: August 29, 2002, 16:51:24
 Job time: 248 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 16:46:56 ; Search time 31.3 Seconds
(without alignments)
2608.741 Million cell updates/sec

Title: US-09-841-758-1
Perfect score: 2550
Sequence: 1 MATKCGNGPGYSTPLEAMK.....PALAHELRYPGDCSSDIWI 472

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_19:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2550	100.0	472	4	Q96GX7
2	2287	89.7	472	11	Q91X87
3	1612.5	63.2	488	10	Q93WS1
4	1587.5	62.3	480	10	Q9LK38
5	1568.5	61.5	487	10	Q93VA1
6	1545.5	60.6	487	10	Q9ASS5
7	1545.5	60.6	487	10	Q93WNO
8	1535.5	60.2	478	10	Q93IH7
9	1524	59.8	457	10	Q9AV46
10	1503	58.9	470	10	Q93WU0
11	1479	58.0	478	10	Q23265
12	1347.5	52.8	471	5	Q9XXF9
13	1247.5	48.9	468	5	Q9VZF4
14	1178	46.2	235	4	Q9H8A8
15	923.5	36.2	462	17	Q976Y0
16	901.5	35.4	466	17	Q97UY0

17	433	17.0	84	11	Q9R0W8	Q9R0W8 mus spretus
18	431	16.9	84	11	Q9R0W7	Q9R0W7 mus spretus
19	393	15.4	146	10	Q42178	Q42178 arabidopsis
20	366	14.4	487	5	O44507	O44507 caenorhabdi
21	250	9.8	68	13	Q9DFK9	Q9DFK9 gillinchlys
22	220.5	8.6	64	10	O42227	O42227 arabidopsis
23	111	4.4	885	3	O07533	O07533 saccharomyc
24	109.5	4.3	764	16	Q9KUB9	Q9KUB9 vibrio chol
25	108.5	4.3	3460	4	P78509	P78509 homo sapien
26	107.5	4.2	546	12	Q913W7	Q913W7 human parvo
27	107.5	4.2	773	12	Q913W8	Q913W8 human parvo
28	106.5	4.2	554	12	Q912B7	Q912B7 human eryth
29	106.5	4.2	781	12	Q912B8	Q912B8 human eryth
30	106	4.2	546	12	Q913X0	Q913X0 human parvo
31	106	4.2	773	12	Q913X1	Q913X1 human parvo
32	105.5	4.1	781	12	P89318	P89318 human parvo
33	105	4.1	604	3	Q9HFG1	Q9HFG1 aspergillus
34	105	4.1	893	16	Q9X1G2	Q9X1G2 thermotoga
35	104.5	4.1	781	12	P89321	P89321 human parvo
36	104	4.1	554	12	Q9PZS9	Q9PZS9 human parvo
37	104	4.1	781	12	Q9PXT0	Q9PXT0 human parvo
38	103	4.0	781	12	P89316	P89316 human parvo
39	102.5	4.0	534	12	Q90201	Q90201 human parvo
40	102.5	4.0	554	12	Q65790	Q65790 human parvo
41	102.5	4.0	554	12	Q9JGP7	Q9JGP7 human parvo
42	102.5	4.0	554	12	Q9WKL9	Q9WKL9 human parvo
43	102.5	4.0	720	10	Q9SDT4	Q9SDT4 arabidopsis
44	102.5	4.0	769	12	Q9PZT4	Q9PZT4 human parvo
45	102.5	4.0	781	12	O85117	O85117 human parvo

ALIGNMENTS

RESULT 1
ID Q96GX7 PRELIMINARY; PRT: 472 AA.
AC Q96GX7;
DT 01-DEC-2001 (TREMUREL. 19, Created)
DT 01-DEC-2001 (TREMUREL. 19, Last sequence update)
DT 01-DEC-2001 (TREMUREL. 19, Last annotation update)
DE SIMILAR TO SELENIUM BINDING PROTEIN 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON ADENOCARCINOMA;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC009084; AAH09084.1; -
SO SEQUENCE 472 AA; 52391 MW; 6DC68F9B45FEC1BC CRC64;

Query Match 100.0%; Score 2550; DB 4; Length 472;
Best Local Similarity 100.0%; Pred. No. 8.9e-215;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MATKCGNGPGYSTPLEAMKGPREEIYVLPCTIYRNTGTAPDYLATVVDPRKSPQYCOVI 60
DB 1 MATKCGNGPGYSTPLEAMKGPREEIYVLPCTIYRNTGTAPDYLATVVDPRKSPQYCOVI 60
QY 61 HRLMPNPKDELHSGNWTSCSCGDSKSKTKVLVPSLISSRIYVVDVSGSEPAAPKLHK 120
DB 61 HRLMPNPKDELHSGNWTSCSCGDSKSKTKVLVPSLISSRIYVVDVSGSEPAAPKLHK 120
QY 121 VIEPKDIAKCEIAFLRHSCLASGEWISLGVKNGKGFVLIDGEPFVGTWERP 180
DB 121 VIEPKDIAKCEIAFLRHSCLASGEWISLGVKNGKGFVLIDGEPFVGTWERP 180
QY 181 GGAAPLGIDFVYQPRHVMVSTEWAAAPNVLDGFNPADVENGLGSHLYVMDQRIETVQ 240
DB 181 GGAAPLGIDFVYQPRHVMVSTEWAAAPNVLDGFNPADVENGLGSHLYVMDQRIETVQ 240

```
Db 181 GGAAPLGDVFWYQPRHNWIMISTEMAAPVNLKRDGENPADVEAGLYSHLYVMDQREHETVQ 240
Oy 241 TLSTKDGLPLEIRFLRHPNDAAGFVGALSTTIORFYKNEGTSVEKVIQVPPKKYK 300
Db 241 TLSTKDGLPLEIRFLRHPNDAAGFVGALSTTIORFYKNEGTSVEKVIQVPPKKYK 300
Oy 301 WLLPMPGLITDILSLDDRFLYFSNMWLGHDROYDISDPORPLTGOLFGLGSIYKGP 360
Db 301 WLLPMPGLITDILSLDDRFLYFSNMWLGHDROYDISDPORPLTGOLFGLGSIYKGP 360
Oy 361 VOVLDEDELSQPEPLVYKGRVAGGPMIOISLDGKRLYITTSIYSAMDKOFYDILIRE 420
Db 361 VOVLDEDELSQPEPLVYKGRVAGGPMIOISLDGKRLYITTSIYSAMDKOFYDILIRE 420
Oy 421 GSVMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472
Db 421 GSVMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472

RESULT 2
Oy 091X87 PRELIMINARY; PRT; 472 AA.
AC 091X87;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE UNKNOWN (PROTEIN FOR MGC:18519).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC011202; AAH11202.1; -
SQ SEQUENCE 472 AA; 52514 MW; 099BCE085668DE36 CRC64;

Query Match 89.7%; Score 2287; DB 11; Length 472;
Best Local Similarity 87.7%; Pred. No. 9,6e-192;
Matches 414; Conservative 27; Mismatches 31; Indels 0; Gaps 0;

Oy 1 MATKCGNGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQYCOVI 60
Db 1 MATKCTKCGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQYCVI 60
Oy 61 HRLPMPNLKDELHSHGWNMTSCSCFSDSTKSRKVLPLSLISSRIYVVDVGSSEPRAPKLHK 120
Db 61 HRLPMPYLKDELHSHGWNMTSCSCFSDSTKSRNKLPLGLISSRIYVVDVGSSEPRAPKLHK 120
Oy 121 VLEPDIHAKCELAFLHTSHCLASGEVMISLGDYKNGKGGFVLLDGETFEVKTWBERP 180
Db 121 VLEPDIHAKCELAFLHTSHCLASGEVMISLGDYKNGKGGFVLLDGETFEVKTWBERP 180
Oy 121 VLEAEIQAQKAVSSLIHSHCLASGEVMVSTIGDLOGNKGFFVLLDGETFEVKTWBERP 180
Db 121 VLEAEIQAQKAVSSLIHSHCLASGEVMVSTIGDLOGNKGFFVLLDGETFEVKTWBERP 180
Oy 181 GGAAPLGDVFWYQPRHNWIMISTEMAAPVNLKRDGENPADVEAGLYSHLYVMDQREHETVQ 240
Db 181 GGAAPMGVDFWYQPRHNWVSTEMAAPVNFKGFNPDAHVEAGLYGSRIFVMDQREHETIQ 240
Oy 241 TLSTKDGLPLEIRFLRHPNDAAGFVGALSTTIORFYKNEGTSVEKVIQVPPKKYK 300
Db 241 TLSTKDGLPLEIRFLRHPNDAAGFVGALSTTIORFYKNEGTSVEKVIQVPPKKYK 300
Oy 301 WLLPMPGLITDILSLDDRFLYFSNMWLGHDROYDISDPORPLTGOLFGLGSIYKGP 360
Db 301 WLLPMPGLITDILSLDDRFLYFSNMWLGHDROYDISDPORPLTGOLFGLGSIYKGP 360
Oy 361 VOVLDEDELSQPEPLVYKGRVAGGPMIOISLDGKRLYITTSIYSAMDKOFYDILIRE 420
Db 361 VOVLDEDELSQPEPLVYKGRVAGGPMIOISLDGKRLYITTSIYSAMDKOFYDILIRE 420
Oy 421 GSVMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472
Db 421 GSVMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472
```

```
Db 421 GSVMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472
Oy 421 GSVMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472

RESULT 3
Oy 093WS1 PRELIMINARY; PRT; 488 AA.
AC 093WS1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE SLENIUM BINDING PROTEIN.
GN SBL.
OS Medicago sativa (Alfalfa).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
OX NCBI_TaxID=3879;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SOMATIC EMBRYOS;
RA Flegmetakis E.;
RT "A novel lotus japonicus nodulin gene is highly conserved among plants
RT and animals, and encodes a homologue to the mammalian selenium binding
RT protein."
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ401228; CAC67501.1; -
SQ SEQUENCE 488 AA; 54109 MW; D0BD0E38D3F50B5D CRC64;

Query Match 63.2%; Score 1612.5; DB 10; Length 488;
Best Local Similarity 62.1%; Pred. No. 1.2e-132;
Matches 292; Conservative 65; Mismatches 110; Indels 3; Gaps 3;

Oy 5 CGNCGPGYSTPLEAMKGPREEIYVLPCTYRNTGTEAPDYLATVDVDPKSPQYCOVIHRLP 64
Db 20 CCKSGPGYASPLEAMSGRREITLYTAVAGTGIEKRPYLATVDLDPSPYSKYIHRLP 79
Oy 65 MPNLKDELHSHGWNMTSCSCFSDSTKSRKVLPLSLISSRIYVVDVGSSEPRAPKLHVIEP 124
Db 80 VPYVDEDLHHTGWNSSCHGDPASQRRFLVPGIVSGRVVYIDTKTPRAPSLHKVVEP 139
Oy 125 KDIAHKLCELAFLHTSHCLASGEVMISLGDYKNGKG-GFVLLDGETFEVKTWBERPGA 183
Db 140 EDISTKGTGLAPRHSHCLASGEIMTSCIGDKDGNAGMGFLDSE-FNVGRMEKPGHS 198
Oy 184 APLGDFWYQPRHNWIMISTEMAAPVNLKRDGENPADVEAGLYSHLYVMDQREHETVQLS 243
Db 199 PLFGYDFWYQPRHNWIMISTEMCAPKAFLOGFRLQHVADGIVGRHLHYVSWFGGELIKOTLD 258
Oy 244 LKD-GLIPLEIRFLRHPNDAAGFVGALSTTIORFYKNEGTSVEKVIQVPPKKYKGL 302
Db 259 LGDKGLPLEIRFLRHPNDAAGFVGALSTTIORFYKNEGTSVEKVIQVPPKKYKGL 318
Oy 303 LPEMGLITDILSLDDRFLYFSNMWLGHDROYDISDPORPLTGOLFGLGSIYKGPVQ 362
Db 319 LPEMGLITDILSLDDRFLYFSNMWLGHDROYDISDPORPLTGOLFGLGSIYKGPVQ 378
Oy 363 VLEDEDELSQPEPLVYKGRVAGGPMIOISLDGKRLYITTSIYSAMDKOFYDILIREGS 422
Db 379 AVKDDGFWQSDVPEIQKRLKRGSPOMIOLSLDGKRLYITTSIYSAMDKOFYDILIREGS 438
Oy 423 VMLQVDVDTYKGGKLKLPNFLVDFGKEPLGAPALAHLELRYPGDCSSDIWI 472
Db 439 HILQIDVDTENGGLKLPNFLVDFGAEEDPSLAHEMRYPGDCSSDIWI 488

RESULT 4
Oy 09LK38 PRELIMINARY; PRT; 480 AA.
AC 09LK38;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
```


DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Selenium-BINDING PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20363099; PubMed=10907853;
 RA Nakamura Y.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
 RT Sequence features of the regions of 4,251,695 bp covered by ninety P1,
 RT TAC and BAC clones.";
 RL DNA Res. 7:217-221(2000).
 DR EMBL: AP000377; BAB01856.1; "-
 DR InterPro: IPR001412; tRNA-synt.1
 DR PROSITE; PS00178; AA_TRNA_LIGASE_I; UNKNOWN.1.
 SQ SEQUENCE 480 AA; 53359 MW; 4508E02A513FA56B CRC64;

Query Match 62.3%; Score 1587.5; DB 10; Length 480;
 Best Local Similarity 61.3%; Pred. No. 1.8e-130;

Matches 289; Conservative 63; Mismatches 116; Indels 3; Gaps 3;

QY 5 CGNGPGYSTPLEAMKGRREIVYLPCTYRNTGTEADPYLATVVDKSPQYCOVIRLP 64
 Db 12 CCKSGPGYAPFLAMSGRRELIYVAAYTGQAKPDYLATVVDPESSSYSSVIRLP 71
 QY 65 MPNKLDELHSGMNTSCSCFGDSTKSRKTLVPSLISRIYVVDGSEPPAKLHKYEP 124
 Db 72 MPYLEDLHSGMNSCSCYGDSCERYLLIPSLISRIYVDTKTPRPSLHKYVDP 131
 QY 125 KDIAKELAFLEHTSHCLASGEVMISSLDYKNGKG-GFVLLDGEFVEYGTWERPGA 183
 Db 132 AEVELEKGLAYPHQPHCLASGDVYLSCLGDEGNAEGSGFLLDSD-FNFKGRWEKDGNS 190
 QY 184 APLCYDFWYOPRHNVMISTEWAAPNVLRDGFNPADVEAGLGSILYWDQORHIVTSL 243
 Db 191 PLCYDFWYOPRHNMTSTSGKAPAFKGFDKLDVSGGLKRLHYNSWPGQELKQLD 250
 QY 244 LKD-GLIPLEIRFLHNPDAOGFVGCALSTIQRFYKNEGSTSVSEKYIOPPKVKGWL 302
 Db 251 LGDTGLPLLEVRFLEHPDKATGFAGCALSTLVRFKNDDETSHEVALISVEPLKVENMI 310
 QY 303 LPEPRLITDILLSDRFLTFYSNWLHGDLROYDISPQRRPLTQGLFSGSYKGPVQ 362
 Db 311 LPEPRLITDILLSDRFLTFYSNWLHGDLROYDISPQRRPLTQGLFSGSYKGPVQ 370
 QY 363 VLEDEELKSOPFLVYVKGKRVAGPQMIQLSLDKRLYITTSLSAMDKQFPLIRGGS 422
 Db 371 ALGEBGKAFQDVPKIRGRLRGPRQFQLSLDKRLYITNSLFSVMDRQFPELVKGS 430
 QY 423 VMLQVDVTVYVGGILKLPNFLVDFGKEPLGALAHLELRYPGDCSDIWI 472
 Db 431 HMLQIDVDTEKGLISINPNFVDGTEPDPGSLAHMKRYPGDCDTSIWI 480

RESULT 5
 ID 093VAL PRELIMINARY; PRT; 487 AA.
 AC 093VAL;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DE Selenium BINDING PROTEIN.
 GN SPP.

OS Lotus japonicus.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
 OX NCBI_TaxID=34305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Flenetakis E.;
 RT "A novel Lotus japonicus nodulin gene is highly conserved among plants
 RT and animals, and encodes a homologue to the mammalian selenium binding
 RT protein.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ401227; CAC67492.1; "-
 DR EMBL: AJ401226; CAC67491.1; "-
 SQ SEQUENCE 487 AA; 53652 MW; DED4C31A908CC12E CRC64;

Query Match 61.5%; Score 1568.5; DB 10; Length 487;
 Best Local Similarity 61.5%; Pred. No. 8.4e-129;

Matches 289; Conservative 63; Mismatches 115; Indels 3; Gaps 3;

QY 5 CGNGPGYSTPLEAMKGRREIVYLPCTYRNTGTEADPYLATVVDKSPQYCOVIRLP 64
 Db 19 CCKSGPGYASPIESMSGRESLIYTAAGTGLEKPDYLATVVDPSPPYSKYVIRLP 78
 QY 65 MPNKLDELHSGMNTSCSCFGDSTKSRKTLVPSLISRIYVVDGSEPPAKLHKYEP 124
 Db 79 VPYVGDELHSGMNSCSCCHDPSAERREFLAPLVSGRITVDTKTPRPSLHKYVDP 138
 QY 125 KDIAKELAFLEHTSHCLASGEVMISSLDYKNGKG-GFVLLDGEFVEYGTWERPGA 183
 Db 139 ADIIOKGLAYPHRTSHCLASGELMISCLGDKGNAEGSGFLLDSD-FNFKGRWEKDGNS 197
 QY 184 APLCYDFWYOPRHNVMISTEWAAPNVLRDGFNPADVEAGLGSILYWDQORHIVTSL 243
 Db 198 PLFYDFWYOPRHNMTSTSGAPAFKGFQDQHVADGLYGRHLTQIQLRGLRQLD 257
 QY 244 L-KDGLIPLEIRFLHNPDAOGFVGCALSTIQRFYKNEGSTSVSEKYIOPPKVKGWL 302
 Db 258 LGNGLIPLEIRFLHNPDAOGFVGCALSTIQRFYKNEGSTSVSEKYIOPPKVKGWL 317
 QY 303 LPEPRLITDILLSDRFLTFYSNWLHGDLROYDISPQRRPLTQGLFSGSYKGPVQ 362
 Db 318 LPEPRLITDILLSDRFLTFYSNWLHGDLROYDISPQRRPLTQGLFSGSYKGPVQ 377
 QY 363 VLEDEELKSOPFLVYVKGKRVAGPQMIQLSLDKRLYITTSLSAMDKQFPLIRGGS 422
 Db 378 ALGEDGKTWSDVPEIQGQKLRGSPQMIQLSLDKRLYITNSLFSVMDRQFPELVKGS 437
 QY 423 VMLQVDVTVYVGGILKLPNFLVDFGKEPLGALAHLELRYPGDCSDIWI 472
 Db 438 HMLQIDVDTEKGLISINPNFVDGTEPDPGSLAHMKRYPGDCDTSIWI 487

RESULT 6
 ID 09ASS5 PRELIMINARY; PRT; 487 AA.
 AC 09ASS5;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DE AT4G14040/DL3061C.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jones T., Kamiya A., Katlin-Neumann G., Kawai J., Lam B.,
 RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,

QY	69	KDELHSGMNWSSCGFSGTSSRFKLYVPLSISSRIYVWVGSPPRPYKLAHYEPDIH	128
		: : : : : : : : : : :	
Db	67	GDELIHSGMNWSSCGSHDASVDKRYLYVPLSISSRIYAITDKEMPRAPSLIKKYDKEITA	126
QY	129	AKCELAFLHTSHCLASGEVMTISSLDGVKNGKG - GVVLLDGETTEFVYGTWEREGGAPLG	187
Db	127	DKTGAFAPHTAHC - RTGEILVSCLDGEBGNAKNGGLLLDSD - FNTKNMEKEDISPLYG	184
QY	188	YDFVYQRPNNVMTENMAAPNVLRDGFNPADVEAGLYGSHLYWDMORHEIYVOTLSL - KD	246
		: : : : : : : : : :	
Db	185	YHFWQRPKHKMTISTWGAPKAFKSGFNLLHYADGLGSHLHYHWSWGGEIKQLIDIGPT	244
QY	247	GLPIREIFLNPDAAGFVSCALSIITQRIYKKEGTSWEVYQVPRKKYVGMILLPEM	306
		: : : : : : : : : :	
Db	245	GLPIREIFLNPDRKDMTGFGSALSSNMIRPKNSDETWSEHAVIISKPLKYVNMILLPEM	304
QY	307	PGLITDILLSLDDEFLFYSNMLHGDLEQYDISDPQREPLTGOGLFGSIYKGGPQVLYED	366
		: : : : : : : : : :	
Db	305	PGLITDPLFSLDDEFLFYVMNLHGDINOYNIEDPKNNVLYLQIYWGSLGKSGPYAVGE	364
QY	367	ELKSQREPLLYKGRVYAGGPQMTQLSLDGKRLYITTSLYSAMDQKQFVPLIREGSMVIO	426
		: : : : : : : : : : :	
Db	365	DGNFTQFEVPPQIKKSLRSGPOMTQLSLDGKRLYATSLSESAWVSPFPEIMKEGSHIQ	424
QY	427	VDVQVYKGLTNPFLYDECKEPLGALAHLELRYPCGDCSSDIWI	472
		: : : : : : : : : : :	
Db	425	IDGTEKGGGLTINPFLYVDEDEPDGSLAHMKMTYPCGGDTSIDIWI	470

Db	79	MPYIGDELHHTGNNSCSCCHDASADRRYLVELLISGRITAYIDTKNDPRAPSLKYVER	138
Qy	125	KDIHAKCLAPLHTSHCLASEVWISSIGDYKNGKG-GEVLLDGEFFEVKGTWERGGA	163
Db	139	KEIEMKGLAPRHTSHCLASGDMVLSCIGDEGNAGKNGEFLDSD-FNYSRWDKGGHG	197
Qy	184	APLGYDPMYQRRHVMSTEMAANVLDRGNPNADVAGLIGSHLYWMDQRIHIVOTLS	243
Db	198	PLFGEYPMYQRFXTMISTSWGAKAPSKGNLOHAVDGLIGSHLHYQPEGMKQITTY	257
Qy	244	LKDGILPELPELFNPDAOGFEVCALSTFORFKNNGEWSVEKYIQVPPKVKQML	303
Db	258	WVFYVMWLQIRFLNDPSKDTGYVSALSSNNIRFKKSDDTWSH-----AIENWIL	309
Qy	304	PEMBGLITDILLSDRFLYFSNMLHGDRLROYDISDOPRPLTGQLELFGSIVYKGFVQY	363
Db	310	PEMBGLITDFLISDDRFEFYVNMHLHGDRIQRYNIEDPKNPVLTOIWMGELLQKSPYKA	369
Qy	364	LEDEBLKSQRPPLVYKKGKRAVGQOMQLSLDGKRLYITTSLYSAMKORYPDLIRGSV	423
Db	370	VGEDEMYQFQVNPQIKKSKSLRAGQOMQLSLDGKRLYATNSLSAMROEYPRTEIMEKSH	429
Qy	424	MLQYDVTVFKGGLKNLNFLEDEKPELGPLALHELHYPGDCSSDITW	472
Db	430	TIQIDVDTDKGGLTLNDFYVDFEDEDGPRALAHENMYPGGDCSTDITW	478

RESULT	11		
023265			
ID	023265	PRELIMINARY;	PRN; 478 AA.
AC	023265;		
DT	01-JAN-1998 (TREMBLrel, 05, Created)		
DT	01-NOV-1999 (TREMBLrel, 12, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel, 19, Last annotation update)		
DE	HYPOTHEtical 53.3 KDA PROTEIN (SELENIUM-BINDING PROTEIN LIKE).		
GN	DL061C OR A14G1040.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.		
OX	NCBI_Taxid=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Beyan M., Stelekema W., Murphy G., Wambutt R., Pohl T., Terryn N.,		
RA	Kreis M., Kavanagh T., Entian K.D., Rieger W., James R.,		
RA	Pulgomenech P., Hatzopoulos P., Obermayer B., Diesterhoft A.,		
RA	Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,		
RA	Schneeller C., Chalwatzis N;		
RL	submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	EU Arabidopsis sequencing project;		
RA	submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; Z97335; CAB4600.1; -;		
DR	EMBL; AL161537; CAB78446.1; -;		
DR	InterPro; IPR002885; PPR.		
DR	InterPro; IPR000504; RRM.		
DR	Pfam; PF01535; PPR; 6.		
DR	PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.		
DR	SEQUENCE 478 AA; 53275 MW; FCE942FEB96D82B CRC64;		

```

RESULT 12
09XXFE9
ID 09XXFE9 PRELIMINARY; PRT; 471 AA.
AC 09XXFE9
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Y37A1B.5 PROTEIN.
GN Y37A1B.5
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxId=6239;
[1]
RN
RP SEQUENCE FROM N.A.
RA McMurtry A.A.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA none;
RX MEDLINE=99069613; PubMed=9851916;
RA "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RT Science 282:2012-2018(1998).
RL EMBL; AL023835; CAAL19490.1; -.
DR InterPro: IPR000345; Cytc_heme_bind.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN.1.
SQ SEQUENCE 471 AA; 52424 MW; C4E95E16E14E1C50 CRC64;
Query Match 52.8%; Score 1347.5; DB 5; Length 471;
Best Local Similarity 52.2%; Pred. No. 1,86-109;
Matches 250; Conservative 85; Mismatches 129; Indels 15; Gaps
8;

```

	Query Match	58.0%;	Score 1479;	DB 10;	length 478;
	Best Local Similarity	57.6%;	Pred. No. 5,6e-121;		
	Matches	270;	Conservative	70;	Mismatches 119; Indels 10; Gaps 3;
QY	5	CGCNGCGCTPLEAMKGRREIVYLPCILRYNTGTAPDTLAVDNDPKSPQCYVITHRP	64		
		: : : : :			
DG	19	CCKSGPGATPLAAAGPREKLITYALTLSGGRRKPDLTAVDDPSSSTVSIVHTRK	78		
LY	65	MPNLDELHHSGWMTCCSCFGDSFRKRLVLPGLISSRIYVDVGSEBRRAKLRKVIEP	124		

[illegible]

RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Yamamoto K.,
RA Arima M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakematsu A., Nakamura Y., Nagahari K., Masuo Y., Oshima A.,
RA "NEO human cDNA sequencing project."
RA Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AK023875; BAB14709.1; -
DR InterPro: IPR000531; TonB_boxC.
DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; UNKNOWN.1.
SQ SEQUENCE 235 AA; 25953 MW; D68FF73A616D986 CRC64;

Query Match 46.2%; Score 1178; DB 4; Length 235;
Best Local Similarity 100.0%; Pred. No. 4.5e-95;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 MNMLDELHSGWNTSCSGFGDSTKSRKTLVPLSLSSRIYVVDGSEPRAKLHVTEP 124
DB 1 MNMLDELHSGWNTSCSGFGDSTKSRKTLVPLSLSSRIYVVDGSEPRAKLHVTEP 60
QY 125 KDIAKCELAFLHTSHCLASGEVMISSLDVKGNGKGFVLLDGETFEVKGWTERPGAA 184
DB 61 KDIAKCELAFLHTSHCLASGEVMISSLDVKGNGKGFVLLDGETFEVKGWTERPGAA 120
QY 185 PLGYDFWQPRHNVMISTEMAAPNVLRDGFNPADYAGLYGSHLYVMDQREHIVOTLSL 244
DB 121 PLGYDFWQPRHNVMISTEMAAPNVLRDGFNPADYAGLYGSHLYVMDQREHIVOTLSL 180
QY 245 KGLIPLERFLHNPDAAGFVGCALSSITQREYKNE 281
DB 181 KGLIPLERFLHNPDAAGFVGCALSSITQREYKNE 217

RESULT 15
0976Y0 PRELIMINARY; PRT: 462 AA.
AC 0976Y0:
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)
DE PUTATIVE SELENIUM-BINDING PROTEIN.
CN ST0059.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N. A.
RC STRAIN-JCM 10545 / 7;
RX PubMed-11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.,
RT "Complete genome sequence of an aerobic thermophilic
RT Crenarchaeon, Sulfolobus tokodaii strain7."
RL DNA Res. 8:123-140(2001).
DR EMBL: AP000981; BAB65016.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 462 AA; 52350 MW; 8AB74CD852CEF7F1 CRC64;

Query Match 36.2%; Score 923.5; DB 17; Length 462;
Best Local Similarity 39.9%; Pred. No. 2.4e-72;
Matches 189; Conservative 93; Mismatches 147; Indels 45; Gaps 12;

QY 12 YSTPLEAMGPREETIVYPCIRYNTGEAPDYATVDVDPKSPQYCOVTHRLPMPNLKDE 71
DB 13 YSPFKAMAKAPEDLAYVACLTGTGINRADFLAVVDVAPKSETYSKIYHAKVELPYINDE 72
QY 72 LHHSGWNTSCSGFGDSTKSRKTLVPLSLSSRIYVVDGSEPRAPKLHVTEPKDIH 128

DB 73 LHHFGWNAACSSALCPNGKPNIERRELLVGLKSSRIYITDTPNPREPITIVIEPEVK 132
QY 129 ACCELAFLHTSHCLASGEVMISSLDVKGNGKGFVLLDGETFEVKGWTERPGAAPIGY 188
DB 133 KVSQSRHLHTVHC-GPDALYISALGNEBEGCGILMDHTSFEPGLKWEIDRGQIYAY 191
QY 189 DFWYQPRHNVMISTEMAAPNVLRDGFNPADYAGLYGSHLYVMDQREHIVOTLSL-KDG 247
DB 192 DFWMNLPLNEVLVSSBMAVNTIEDGLKLEHLK-DRYGNRIHFWDLRKRRIHSLTLEEN 250
QY 248 LPLERFLHNPDAAGFVGCALSSITQREYKNEGVSEKTVQVPPKKVKGKL 302
DB 251 RMALELRPLHDTKLMGFTNMVYSLKDLSSSIWLMF-YEDGKMAEKVLEIPEPLUG-N 308
QY 303 LPE-----MPGLITDILLSDREFLYFSNMLHGDLRQYDISDPQRPRLTGQLPLGSI 355
DB 309 LPEILKPFKAVPPLVTDIDISLDKFLYLSMGIGEVROYDISNPFKPLVLTGKVLGTF 368
QY 356 VKGFPVQVLEDELKSQPEPLVYKGRVAGSPQMIQLSDGRRLYITTSYAMDKQYTP 415
DB 369 HRA-----DHP-----AGHKLIGAPQMLEISHDGRVYVYVNSLYSTWNOQFYP 411
QY 416 DLIRGSAVLQVDVDTYVGGTLKLNPNELVDREKKEPLGPAALAEHLRYPGSDCSSD 469
DB 412 EELKGMVAKLANAN---PSGGLTEIDKEFFVDFGE-----ARSHQVRLSGDASSD 457

Search completed: August 29, 2002, 16:51:03
Job time: 247 sec

THIS PAGE BLANK (USPTO)



Creation date: 01-16-2004
Indexing Officer: TLAM2 - THY LAM
Team: OIPEBackFileIndexing
Dossier: 09841758

Legal Date: 12-23-2002

No.	Doccode	Number of pages
1	NPL	6 ✓
2	NPL	38 ✓
3	NPL	8 ✓
4	NPL	6 ✓
5	NPL	24 ✓
6	NPL	6 ✓
7	NPL	32 ✓
8	NPL	10
9	NPL	12

Total number of pages: 142

Remarks:

Order of re-scan issued on

